## GenCore version 5.1.6 Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2005, 15:38:47; Search time 8644.89 Seconds

(without alignments)

12117.322 Million cell updates/sec

Title: US-10-624-932-1

Perfect score: 2752

Sequence: 1 ccgcggggccccgcgcccgg.....tgagtgctgaggccggccag 2752

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 segs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_htc:\*

4: gb est3:\*

5: gb est4:\*

6: gb est5:\*

7: gb\_est6:\*

8: gb gss1:\*

9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

|   |       |       | *     |        |    |          |                    |
|---|-------|-------|-------|--------|----|----------|--------------------|
| R | esult |       | Query |        |    |          |                    |
|   | No.   | Score | Match | Length | DB | ID       | Description        |
|   |       |       |       |        |    |          | 24406401 4         |
|   | Ţ     | 960.2 | 34.9  | 2802   | 9  | AY406491 | AY406491 Homo sapi |
|   | 2     | 950.4 | 34.5  | 2791   | 9  | AY406493 | AY406493 Mus muscu |
|   | 3     | 923.4 | 33.6  | 3790   | 3  | AK031655 | AK031655 Mus muscu |
|   | 4     | 884   | 32.1  | 1852   | 3  | CR598115 | CR598115 full-leng |
|   | 5     | 872.4 | 31.7  | 3866   | 3  | AK018177 | AK018177 Mus muscu |
|   | 6     | 814   | 29.6  | 2802   | 9  | AY406492 | AY406492 Pan trogl |
|   | 7     | 810.4 | 29.4  | 2532   | 9  | AY411747 | AY411747 Homo sapi |
|   | 8     | 780.4 | 28.4  | 2532   | 9  | AY411749 | AY411749 Mus muscu |
|   |       |       |       |        |    |          |                    |

|   | 9  | 768.6 | 27.9 | 1034  | 4 | BI758231 | BI758231 6 | 03029876  |
|---|----|-------|------|-------|---|----------|------------|-----------|
|   | 10 | 736.6 | 26.8 | 1532  | 3 | BC033727 | ВС033727 Н | Iomo sapi |
|   | 11 | 735.2 | 26.7 | 788   | 1 | AI951556 | AI951556 w | rv36f04.x |
| : | 12 | 721.6 | 26.2 | 796   | 5 | BX348193 | BX348193 B | 3X348193  |
|   | 13 | 716.4 | 26.0 | 818   | 4 | BI818609 | BI818609 6 | 03033362  |
|   | 14 | 692   | 25.1 | 853   | 5 | BX364574 | BX364574 B | 3X364574  |
|   | 15 | 678.2 | 24.6 | 2532  | 9 | AY411748 | AY411748 P | an trogl  |
|   | 16 | 672.6 | 24.4 | 934   | 2 | BF311804 | BF311804 6 | 01897316  |
|   | 17 | 666   | 24.2 | 900   | 5 | BX345406 | BX345406 B | 3X345406  |
|   | 18 | 665   | 24.2 | 859   | 2 | BF311896 | BF311896 6 | 01897733  |
|   | 19 | 645.2 | 23.4 | . 977 | 5 | BX345407 | BX345407 B | 3X345407  |
|   | 20 | 631.2 | 22.9 | 1072  | 5 | BX422753 | BX422753 B | 3X422753  |
|   | 21 | 618.8 | 22.5 | 756   | 5 | BU612387 | BU612387 U | JI-M-EWO- |
|   | 22 | 613.2 | 22.3 | 874   | 5 | BQ689148 | BQ689148 A | GENCOURT  |
| • | 23 | 610.2 | 22.2 | 2775  | 9 | AY401471 | AY401471 M | lus muscu |
|   | 24 | 607   | 22.1 | 889   | 5 | BQ691915 | BQ691915 A | GENCOURT  |
|   | 25 | 599   | 21.8 | 2775  | 9 | AY401469 | AY401469 H | iomo sapi |
|   | 26 | 595.2 | 21.6 | 601   | 1 | AL516580 | AL516580 A |           |
|   | 27 | 589.4 | 21.4 | 604   | 2 | BE314370 | BE314370 6 |           |
|   | 28 | 566.6 | 20.6 | 678   | 6 | CA749784 | CA749784 U |           |
|   | 29 | 561   | 20.4 | 1175  | 2 | BF530640 | BF530640 6 |           |
|   | 30 | 555.6 | 20.2 | 788   | 6 | CA317532 | CA317532 U |           |
|   | 31 | 555   | 20.2 | 572   | 7 | CR554569 | CR554569 D |           |
|   | 32 | 494.2 | 18.0 | 499   | 7 | CR747398 | CR747398 C |           |
|   | 33 | 486.2 | 17.7 | 548   | 5 | BX452510 | BX452510 B |           |
|   | 34 | 475.4 | 17.3 | 824   | 4 | BI737024 | BI737024 6 |           |
|   | 35 | 471.2 | 17.1 | 540   | 4 | BI467799 | BI467799 3 |           |
|   | 36 | 470.8 | 17.1 | 2507  | 9 | AY401470 | AY401470 P | _         |
|   | 37 | 438.2 | 15.9 | 675   | 6 | CA315487 | CA315487 U |           |
|   | 38 | 433   | 15.7 | 471   | 5 | BX282095 | BX282095 B |           |
|   | 39 | 428.4 | 15.6 | 1147  | 5 | BU840446 | BU840446 A |           |
|   | 40 | 424   | 15.4 | 843   | 4 | BG913440 | BG913440 6 |           |
|   | 41 | 417.8 | 15.2 | 856   | 7 | CN164143 | CN164143 9 |           |
|   | 42 | 411.8 | 15.0 | 749   | 7 | CF735417 | CF735417 U |           |
|   | 43 | 410.6 | 14.9 | 1111  | 4 | BG298307 | BG298307 6 |           |
|   | 44 | 409.8 | 14.9 | 751   | 7 | CF735550 | CF735550 U |           |
|   | 45 | 408.8 | 14.9 | 460   | 2 | BF443156 | BF443156 2 | 260542 MA |

С

RESULT 1

## ALIGNMENTS

```
AY406491
                                                                 GSS 15-DEC-2003
                                     2802 bp
                                                        linear
LOCUS
                                                DNA
DEFINITION
            Homo sapiens UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION
            AY406491
VERSION
            AY406491.1 GI:39762465
KEYWORDS
            GSS.
            Homo sapiens (human)
SOURCE
            Homo sapiens
  ORGANISM
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
            1 (bases 1 to 2802)
  AUTHORS
            Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
            Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
```

```
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
          Adams, M.D. and Cargill, M.
          Inferring nonneutral evolution from human-chimp-mouse orthologous
 TITLE
          gene trios
          Science 302 (5652), 1960-1963 (2003)
 JOURNAL
  PUBMED
          14671302
          2 (bases 1 to 2802)
REFERENCE
          Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 AUTHORS
          Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
          Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
          Adams, M.D. and Cargill, M.
 TITLE
          Direct Submission
          Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 JOURNAL
          Rockville, MD 20850, USA
          This sequence was made by sequencing genomic exons and ordering
COMMENT
          them based on alignment.
                   Location/Qualifiers
FEATURES
                   1. .2802
    source
                   /organism="Homo sapiens"
                   /mol type="genomic DNA"
                   /db xref="taxon:9606"
                   <1. .>2802
    gene
                   /gene="UNC5C"
                   /locus tag="HCM2575"
ORIGIN
                       34.9%; Score 960.2; DB 9;
                                                 Length 2802;
 Query Match
 Best Local Similarity
                       61.3%; Pred. No. 6.4e-192;
 Matches 1682; Conservative
                             0; Mismatches 995;
                                                          69;
                                                              Gaps
                                                                      6;
                                                 Indels
          57 GCCCGGCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGG 116
Qу
                            \perp
                         1
          66 GCTCGTGCTACCTGCCCTGGCCCTGCTCAGCGCCAGCGGCACTGGCTCCGCCGCCCAAGA 125
Db
         117 TGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCC 176
Qу
                                I I I I I
                                                              -11 -11
         126 TGATGACTTTTTCATGAACTCCCAGAAACTTTTCCTTCTGATCCACCTGAGCCTCTGCC 185
Db
         177 CCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGT 236
Qу
             186 ACATTTCCTTATTGAGCCTGAAGAAGCTTATATTGTGAAGAATAAGCCTGTGAACCTGTA 245
Db
         237 GTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCA 296
Qу
                        246 CTGTAAAGCAAGCCCTGCCACCCAGATCTATTTCAAGTGTAATAGTGAATGGGTTCATCA 305
Db
         297 GGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGT 356
Qу
                                111 1 1
                                                 11 11
                                                          . | ||| ||
               306 GAAGGACCACATAGTAGATGAAAGAGTAGATGAAACTTCCGGTCTCATTGTCCGGGAAGT 365
Db
         357 CCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTG 416
Qу
              366 GAGCATTGAGATTTCGCGCCAGCAAGTGGAAGAACTCTTTGGACCTGAAGATTACTGGTG 425
Db
         417 CCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCAT 476
QУ
             426 CCAGTGTGTGGCCTGGAGCTCCGCGGGTACCACAAAGAGCCGGAAGGCGTATGTGCGCAT 485
Db
```

|   | QУ   | 477  | AGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCT       | 530  |
|---|------|------|--|------|
| • | Db   | 486  | TGCATNNNNNNNNNNCGGAAGACATTTGAGCAGGAACCCCTAGGAAAGGAA          | 545  |
|   | Qу   | 531  | GGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGA | 590  |
|   | Db   | 546  | GGAACAGGAAGTCTTACTCCAGTGTCGACCACCTGAAGGGATCCCAGTGGCTGAGGTGGA | 605  |
|   | Qy   | 591  | GTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCG | 650  |
|   | Db . | 606  | ATGGTTGAAAAATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTTATATTACTAT | 665  |
|   | Qу   | 651  | GGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGT | 710  |
|   | Db   | 666  | TGATCACAACCTCATCATAAAGCAGGCCCGACTCTCTGATACTGCAAATTACACCTGTGT | 725  |
|   | Qy   | 711  | GGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAA    | 770  |
| • | Db   | 726  | TGCCAAAAACATTGTTGCCAAGAGGAAAAGTACAACTGCCACTGTCATAGTCTATGTCAA | 785  |
|   | QУ   | 771  | CGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTG | 830  |
|   | Db   | 786  | CGGTGGCTGGTCCACCTGGACGGAGTGGTCTGTGTAACAGCCGCTGTGGACGAGGGTA   | 845  |
|   | QУ   | 831  | GCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGA | 890  |
|   | Db   | 846  | TCAGAAACGTACAAGGACTTGTACCAACCCGGCACCACTCAATGGGGGTGCCTTCTGTGA | 905  |
|   | Qy   | 891  | GGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAG | 950  |
|   | Db   | 906  | AGGGCAGAGTGTGCAGAAAATAGCCTGTACTACGTTATGCCCAGTGGATGGCAGGTGGAC | 965  |
|   | QУ   | 951  | CCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTG | 1010 |
|   | Db   | 966  | GCCATGGAGCAAGTGGTCTACTTGTGGAACTGAGTGCACCCACTGGCGCAGGAGGGAG   | 1025 |
|   | Qy   |      | CTCTGACCCAGCACCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCG  | 1070 |
| • | Db   |      | CACGGCGCCAGGAATGGAGGCAAGGACTGCGACGGCCTCGTCTTGCAATCCAA        | 1085 |
|   | Qy   | 1071 | CAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTA | 1130 |
|   | Db   | 1086 | GAACTGCACTGATGGGCTTTGCATGCAGACTGCTCCTGATTCAGATGATGTTGCTCTCTA | 1145 |
|   | Qу   | 1131 | TGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTGCTCATCCT       | 1187 |
|   | Db   | 1146 | TGTTGGGATTGTGATAGCAGTGATCGTTTGCCTGGCGATCTCTGTAGTTGTGGCCTTGTT | 1205 |
|   | Qу   | 1188 | CGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCAC | 1247 |
|   | Db   | 1206 | TGTGTATCGGAAGAATCATCGTGACTTTGAGTCAGATATTATTGACTCTTCGGCACTCAA | 1265 |
|   | Qу   | 1248 | CTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCAC | 1307 |
|   | Db   | 1266 |  | 1316 |
|   |      |      |  |      |

.

|     | Qy<br>Db |      | CATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCA    |       |
|-----|----------|------|--|-------|
|     | Qy       | 1368 | GGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCC          | 1418  |
|     | Db       | 1377 |  | 1436  |
|     | Qy       | 1419 | CCTGGGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTT    | 1478  |
|     | Db       | 1437 |  | 1496  |
|     | Qy       | 1479 | CGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACAT | 1538  |
| •   | Db       | 1497 | TACGTCCAAGCTGTCCCCTCAGATGACCCAGTCGTTGTTGGAGAATGAAGCCCTCAGCCT | 1556  |
|     | Qy       | 1539 | GACCTATGGGACCTTCAA   | 1556  |
|     | Db       | 1557 | GAAGAACCAGAGTCTAGCAAGGCAGACTGATCCATCCTGTACCGCATTTGGCAGCTTCAA | 1616` |
|     | Qy       | 1557 | CTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGA  | 1616  |
|     | Db       |      | CTCGCTGGGAGGTCACCTTATTGTTCCCAATTCAGGAGTCAGCTTGCTGATTCCCGCTGG |       |
|     | Qу       | 1617 | TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGT | 1676  |
|     | Db       | 1677 | GGCCATTCCCCAAGGGAGAGTCTACGAAATGTATGTGACTGTACACAGGAAAGAAA     | 1736  |
|     | Qy       | 1677 | GAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC | 1736  |
|     | Db       | 1737 | GAGGCCACCCATGGATGACTCTCAGACACTTTTGACCCCTGTGGTGAGCTGTGGGCCCCC | 1796  |
|     | Qy       |      | TGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCC |       |
|     | Db       | 1797 | AGGAGCTCTGCTCACCCGCCCAGTCGTCCTCACTATGCATCACTGCGCAGACCCCAATAC | 1856  |
|     | Qy       | 1797 | TGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCT | 1856  |
| 0.0 | Db       | 1857 | CGAGGACTGGAAAATACTGCTCAAGAACCAGGCAGCACAGGGACAGTGGGAGGATGTGGT | 1916  |
|     | Qy       | 1857 | GCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTG | 1916  |
|     | Db       | 1917 | GGTGGTCGGGGAGAAACTTCACCACCCCTGCTACATTCAGCTGGATGCAGAGGCCTG    | 1976  |
|     | Qу       | 1917 | CTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGC | 1976  |
|     | Db       | 1977 | CCACATCCTCACAGAGAACCTCAGCACCTACGCCCTGGTAGGACATTCCACCACCAAAGC | 2036  |
|     | Qy       | 1977 | TGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC      | 2036  |
|     | Db       | 2037 | GGCTGCGAAGCGCCTCAAGCTGGCCATCTTTGGGCCCCTGTGCTGCTCCTCGCTGGAGTA | 2096  |
|     | Qу       | 2037 | CAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCT | 2096  |
|     | Db       | 2097 | CAGCATCCGAGTCTACTGTCTGGATGACACCCAGGATGCCCTGAAGGAAATTTTACATCT | 2156  |
|     | Ov       | 2097 | GGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAG | 2156  |

| Db                   | 2157 |  |
|----------------------|------|--|
| Qy                   | 2157 | TTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCT 2216  |
| Db                   | 2217 |  |
| Qу                   | 2217 | CCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTT 2276  |
| Db                   | 2277 | GCTGGCTAAATATCAGGAAATTCCATTTTACCATGTTTGGAGTGGATCTCAAAGAAACCT 2336  |
| Qy                   | 2277 | GCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCT 2336  |
| Db                   | 2337 | GCACTGCACCTTCACTCTGGAAAGATTTAGCCTGAACACAGTGGAGCTGGTTTGCAAACT 2396  |
| Qy                   | 2337 | GTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAA 2396  |
| Db                   | 2397 | CTGTGTGCGGCAGGTGGAAGGAGAGGGCAGATCTTCCAGCTCAACTGCACCGTGTCAGA 2456   |
| Qy                   | 2397 | GGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGT 2456  |
| Db                   | 2457 | GGAACCTACTGGCATCGATTTGCCGCTGCTGGATCCTGCGAACACCATCACCACGGTCAC 2516  |
| Qу                   | 2457 | GGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGA 2516  |
| Db                   | 2517 | GGGGCCCAGTGCTTTCAGCATCCCTCTCCCTATCCGGCAGAAGCTCTGTAGCAGCCTGGA 2576  |
| Qy                   | 2517 | CCCACCTGTAGGCGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAG 2576  |
| Db                   | 2577 | TGCCCCCAGACGAGAGGCCATGACTGGAGGATGCTGGCCCATAAGCTGAACCTGGACAG 2636   |
| Qy                   | 2577 | CCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGGGA 2636   |
| Db                   | 2637 | GTACTTGAATTACTTTGCCACCAAATCCAGCCCAACTGGCGTAATCCTGGATCTTTGGGA 2696  |
| Qy                   | 2637 | GGCGCGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTG   |
| Db                   |      | AGCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGGCAGCTGTCTTGGAAGAAATGGG 2756  |
| Qy                   | 2697 | CCAGCCAGACGCTGGCCTCTCACAGTGTCGGAGGCTGAGTGCTGA 2742   |
| Db                   | 2757 | AAGACATGAAACGGTGGTCCTTAGCAGCAGAAGGGCAGTATTAA 2802  |
| RESULT 2<br>AY406493 |      |  |
| LOCUS<br>DEFINITION  | ON M | Y406493 2791 bp DNA linear GSS 15-DEC-2003 us musculus UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence, |
| ACCESSIO             | -    | enomic survey sequence.<br>Y406493   |
| VERSION<br>KEYWORDS  |      | Y406493.1 GI:39762467<br>SS.   |
| SOURCE               | M    | us musculus (house mouse)  |
| ORGANI               | SM M | us musculus  |

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 2791)

```
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 AUTHORS
           Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
           Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
           Adams, M.D. and Cargill, M.
 TITLE
           Inferring nonneutral evolution from human-chimp-mouse orthologous
           gene trios
           Science 302 (5652), 1960-1963 (2003)
 JOURNAL
  PUBMED
           14671302
             (bases 1 to 2791)
REFERENCE
 AUTHORS
           Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
           Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
           Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
           Adams, M.D. and Cargill, M.
           Direct Submission
 TITLE
 JOURNAL
           Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
           Rockville, MD 20850, USA
           This sequence was made by sequencing genomic exons and ordering
COMMENT
           them based on alignment.
                   Location/Oualifiers
FEATURES
                   1. .2791
    source
                   /organism="Mus musculus"
                   /mol type="genomic DNA"
                   /db xref="taxon:10090"
                   <1. .>2791
    gene
                   /gene="UNC5C"
                   /locus tag="HCM2575"
ORIGIN
                        34.5%;
                               Score 950.4; DB 9;
                                                   Length 2791;
 Query Match
 Best Local Similarity
                        62.3%; Pred. No. 7.5e-190;
                                                            80;
                                                                Gaps
                                                                        7;
 Matches 1653; Conservative
                               0; Mismatches 921;
                                                   Indels
         149 TGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACA 208
Qу
                         \square
         158 TTCCATCTGACCCACCTGAGCCATTGCCACACTTCCTCATTGAGCCCGAGGAAGCTTACA 217
Db
         209 TCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCT 268
Qy
             1 11 11:1111111 111
                                  11
                                        11 11 111
                                                   Db
         269 TCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACG 328
Qy
             278 TCAAGTGCAACAGCGAGTGGGTTCATCAGAAGGACCACGTAGTAGACGAGAGAGTAGATG 337
Db
         329 GGAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGA 388
Qу
               1 \cdot 1 \cdot 1 \cdot 1 \cdot 1 \cdot 1 \cdot 1
                                1 11 11 11:11 1 1 11:1 1 11:11:11 11:1
Db
         338 AAACCTCTGGTCTAATTGTGAGAGAGTGAGCATTGAGATTTCACGCCAGCAGGTGGAGG 397
         389 AGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCA 448
Qу
                         398 AACTGTTTGGGCCTGAAGATTACTGGTGCCAGTGTGTGGCCTGGAGCTCAGCAGGCACTA 457
Db
         449 CCAAGAGTCAGAAGGCCTACATCCGCATAGCCAG----ATTGCGCAAGAACTTCGAGC 502
Qу
             1 111111 11111 11111
                                                    458 CGAAGAGTCGGAAGGCATACGTGCGCATTGCGTTTCCAGATCTGCGGAAGACATTCGAGC 517
Db
Qy
         503 AGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCAC 562
```

| Db     |  |
|--------|--|
| Qy     | 563 CGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGT 622   |
| Db     |  |
| Qу     | 623 CCCTGGACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCC 682   |
| Db     | 627 CTGAAGATCGGAACTTTTATATTACTATCGATCACAACCTGATCATCAAGCAAG             |
| Qy     | 683 TTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCG 742   |
| Db     | 687 TCTCAGATACAGCAAATTATACCTGTGTTGCCAAAAATATTGTTGCCAAGAGAAAAAAGCA 746  |
| Qy     | 743 CCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGTCGACGTGGACCGAGTGGTCCG 802       |
| Db     | 747 CCACAGCCACTGTCATCGTGTATGTTAATGGTGGCTGGTCCACCTGGACAGAGTGGTCTG 806   |
| Qу     | 803 TCTGCAGCGCCAGCTGTGGGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGG 862     |
| Db     | 807 TGTGTAACAGCCGCTGTGGGCGAGGATATCAGAAACGCACAAGAACCTGCACCAACCCAG 866   |
| Qy     | 863 CGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCA 922   |
| Db     | 867 CCCCACTCAATGGTGGGGCCTTCTGTGAGGGGCAGAGTGTGCAGAAAATAGCATGCACTA 926   |
| Qy     | 923 CCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGG 982   |
| Db     | 927 CGTTATGTCCAGTGGATGGTAGGTGGACTTCATGGAGCAAATGGTCAACCTGTGGGACTG 986   |
| Qy     | 983 ACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGG 1042  |
| Db     | 987 AATGCACCCACTGGCGCAGGAGGGGTTACAGCACCAGCCCCCAAGAACGGGGGTAAGG 1046    |
| Qy     | 1043 AGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTG 1102 |
| Db     | 1047 ACTGTGATGGCCTGGTCCTCCAATCCAAGAACTGCACTGATGGGCTGTGCATGCA           |
| Qy     | 1103 CTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCC 1159    |
| . Db   | 1107 CTCCTGACTCAGATGATGTGGCTCTCTACGTGGGGATTGTGATCGCTGTAACAGTCTGTC 1166 |
| Qy     | 1160 TGGTCCTGCTGCTGCTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACT 1219    |
|        | 1167 TGGCGATCACTGTTGTGTGTGTGTGTGTGTGTGTGTGTGT                          |
| Qy<br> | 1220 CAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCA 1279 |
| Db     | 1227 CTGACATCATTGACTCCTCAGCACTCAATGGCGGCTTTCAGCCTGTGAACATCAAG 1282     |
| Qy     | 1280 GCAAAGCAGACACCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCA 1339   |
| Db     | 1283GCTGCCAGACAAGATCTCCTGGCTGTCCCCCCTGACCTCACCTCAGCTGCAGCCA 1337       |
| Qу     | 1340 CCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCT 1394      |

| Db   | 1338 | TGTACAGGGGACCTGTCTATGCTCTGCATGATGTCTCAGACAAAATCCCAATGACCAACT  | 1397 |
|------|------|---|------|
| Qy   | 1395 | CACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTC  | 1454 |
| Db   | 1398 | CTCCAATTCTGGACCCACTACCCAACTTGAAAATCAAAGTGTACAACAGCTCAGGTGCTG  | 1457 |
| Qу   | 1455 | TCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCC                  | 1498 |
| Db   | 1458 | TCACTCCTCAGGATGACCTTGCCGAGTTCTCATCCAAACTGTCACCCCAGATGACCCAGT  | 1517 |
| Qy   | 1499 | AGAACTACTTCCGCTCCCTGCCCCGAGGCA                                | 1528 |
| Db . | 1518 | CCTTGCTAGAGAATGAGGCCCTTAACCTGAAGAACCAGAGCCTCGCAAGACAGAC       | 1577 |
| Qу   | 1529 | CCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATA  | 1588 |
| Db   | 1578 | CATCCTGCACAGCATTTGGTACCTTCAACTCTCTTGGGGGTCACCTCATCATTCCTAATT  | 1637 |
| Qу   | 1589 | CAGGTATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCT   | 1648 |
| Db   | 1638 | CAGGAGTAAGCTTGCTGATTCCCGCTGGGGCCCATTCCTCAGGGGAGAGTCTATGAAATGT | 1697 |
| Qу   | 1649 | ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTG             | 1708 |
| Db   | 1698 | ATGTGACTGTACACAGGAAAAATATGAGGCCCCCCATGGAAGACTCTCAGACCCTAC     | 1757 |
| Qу   | 1709 | TGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGG  | 1768 |
| Db   | 1758 | TTACCCCTGTGGTGAGCTGTGGGCCTCCTGGAGCTCTGCTGACCCGCCCTGTCATCCTCA  | 1817 |
| Qу   | 1769 | CTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGT  | 1828 |
| Db   | 1818 | CTCTGCATCACTGTGCAGACCCCAGCACCGAGGACTGGAAGATCCAGCTCAAAAACCAGG  | 1877 |
| Qу   | 1829 | CGTGCGAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCT  | 1888 |
| Db   | 1878 | CAGTGCAGGGACAATGGGAGGATGTTGTGGTGGTTGGGGAGGAGAACTTCACAACCCCCT  | 1937 |
| Qу   | 1889 | ACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTG  | 1948 |
| Db   | 1938 | GTTACATTCAGCTGGATGCAGAGGCTTGCCATATCCTCACAGAGAACCTCAGTACCTATG  | 1997 |
| Qу   | 1949 | CCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTG  | 2008 |
| Db   | 1998 | CCCTGGTTGGGCAGTCCACCACAAAGCAGCTGCCAAGCGTCTTAAACTGGCCATCTTTG   | 2057 |
| Qу   | 2009 | CGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC           | 2068 |
| Db   | 2058 | GGCCCTCTGCTGCTCTTCCCTGGAGTACAGCATTAGAGTCTACTGCCTGGATGACACAC   | 2117 |
| Qу   | 2069 | ACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGG | 2128 |
| Db   | 2118 | AGGATGCCCTGAAGGAAGTTCTACAACTGGAGAGGCAAATGGGAGGACAGCTCCTAGAAG  | 2177 |
| Qу   | 2129 | AGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC        | 2188 |
| Db   | 2178 | AACCCAAGGCTCTTCATTTTAAAGGCAGCATCCACAACCTGCGCCTGTCTATTCATGACA  | 2237 |

| Qy   |     | 9 TGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATC 2248  |
|--|-----|--|
| Db   | 223 | 8 TCGCCCATTCCCTCTGGAAGAGCAAATTGCTGGCTAAGTATCAGGAAATTCCATTTTACC 2297  |
| Qу   | 224 | 9 ACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCC 2308  |
| Db   | 229 | 8 ACATCTGGAGTGGCTCTCAAAGAAACCTCCACTGCACCTTCACTCTGGAAAGACTCAGCC 2357  |
| Qу   | 230 | 9 CCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGA 2368  |
| Db   | 235 | 8 TAAACACAGTGGAACTGGTTTGCAAACTCTGTGTGCGGCAGGTTGAAGGAGAAGGGCAGA 2417  |
| Qу   | 236 | 9 GCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGG 2428  |
| Db   | 241 | 8 TCTTCCAGCTCAACTGTACTGTGTCAGAGGAACCTACTGGCATCGACTTACCTCTCCTGG 2477  |
| Qy .   | 242 | 9 AGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCA 2488  |
| Db   | 247 | 8 ACCCTGCTAGTACCATCACCACTGTCACCGGACCAAGTGCTTTCAGCATTCCTCTCCCTA 2537  |
| Qy   | 248 | 9 TTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGA 2548  |
| Db   | 253 | 8 TCCGGCAGAAGCTATGCAGCAGCCTGGATGCCCCTCAAACAAGAGGCCATGACTGGAGGA 2597  |
| Qу   | 254 | 9 CTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCC 2608  |
| Db   | 259 | 8 TGCTGGCCCATAAACTCAACCTGGACAGGTACTTGAATTACTTTGCCACCAAATCGAGCC 2657  |
| Qу   | 260 | 9 CCACAGCCATGATCCTCAACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCC 2668   |
| Db   | 265 | 8 CAACTGGCGTAATCCTGGATCTTTGGGAAGCACAGAACTTCCCAGATGGAAACCTGAGCA 2717  |
| Qy   | 266 | 9 AGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGG 2728  |
| Db   | 271 | 8 TGCTGGCAGCCGTCCTGGAAGAAATGGGAAGACATGAGACAGTGGTGTCCTTGGCAGCAG 2777  |
| Qу   | 272 | 9 AGGCTGAGTGCTGA 2742  |
| Db ·   | 277 | 8 AAGGACAGTATTGA 2791  |
| RESULT 3<br>AK031655<br>LOCUS<br>DEFINITI              | (   | AK031655 3790 bp mRNA linear HTC 03-APR-2004<br>Mus musculus 13 days embryo male testis cDNA, RIKEN full-length                      |
|  |     | enriched library, clone:6030473H24 product:unc5 homolog (C. elegans) 3, full insert sequence.  |
| ACCESSION<br>VERSION<br>KEYWORDS<br>SOURCE<br>ORGANISM |     | AK031655 AK031655.1 GI:26327502 HTC; CAP trapper. Mus musculus (house mouse) Mus musculus  |
| REFERENC   | E   | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. |
| AUTHOR   | S   | Carninci, P. and Hayashizaki, Y.   |

```
High-efficiency full-length cDNA cloning
  TITLE
            Meth. Enzymol. 303, 19-44 (1999)
  JOURNAL
 MEDLINE
            99279253
            10349636
   PUBMED
REFERENCE
            2
            Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 AUTHORS
            Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
  TITLE
            prepare full-length cDNA libraries for rapid discovery of new genes
            Genome Res. 10 (10), 1617-1630 (2000)
  JOURNAL
            20499374
  MEDLINE
            11042159
   PUBMED
REFERENCE
            3
            Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
  AUTHORS
            Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
            Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
            Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
            Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
            Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
            Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
            RIKEN integrated sequence analysis (RISA) system--384-format
  TITLE
            sequencing pipeline with 384 multicapillary sequencer
            Genome Res. 10 (11), 1757-1771 (2000)
  JOURNAL
  MEDLINE
            20530913
   PUBMED
            11076861
REFERENCE
            The RIKEN Genome Exploration Research Group Phase II Team and the
  AUTHORS
            FANTOM Consortium.
            Functional annotation of a full-length mouse cDNA collection
  TITLE
  JOURNAL
            Nature 409, 685-690 (2001)
REFERENCE
            5
            The FANTOM Consortium and the RIKEN Genome Exploration Research
  AUTHORS
            Group Phase I & II Team.
  TITLE
            Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
            Nature 420, 563-573 (2002)
  JOURNAL
            6 (bases 1 to 3790)
REFERENCE
            Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
  AUTHORS
            Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
            Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
            Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
            Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
            Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
            Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
            Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
            Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
            Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
            Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
            Muramatsu, M. and Hayashizaki, Y.
            Direct Submission
  TITLE
            Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
  JOURNAL
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
            URL: http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
            Fax: 81-45-503-9216)
```

```
COMMENT
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
           prepare mouse tissues.
            Please visit our web site for further details.
            URL:http://genome.gsc.riken.jp/
           URL:http://fantom.gsc.riken.jp/.
FEATURES
                    Location/Qualifiers
                    1. .3790
     source
                    /organism="Mus musculus"
                    /mol type="mRNA"
                    /strain="C57BL/6J"
                    /db xref="FANTOM DB:6030473H24"
                    /db xref="taxon:10090"
                    /clone="6030473H24"
                    /sex="male"
                    /tissue type="testis"
                    /clone lib="RIKEN full-length enriched mouse cDNA library"
                    /dev stage="13 days embryo"
     CDS ·
                    118. .2970
                    /note="unnamed protein product; putative
                    unc5 homolog (C. elegans) 3 (MGD|MGI:1095412,
                    GB|NM 009472, evidence: BLASTN, 99%, match=464)"
                    /codon start=1
                    /protein id="BAC27495.1"
                    /db xref="GI:26327503"
                    /translation="MRKGLRATAARCGLGIGYLLQMLVLPALALLSASGTGSAAQDDE
                    FFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQ
                    KDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYV
                    RIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEDIIDPAEDRNFYI
                    TIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRC
                    GRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACTTLCPVDGRWTSWSKWSTCGTECTH
                    WRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQGFIYPISTEHRPQNEYGFSSAPD
                    SDDVALYVGIVIAVTVCLAITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAA
                    \mathtt{RQDLLAVPPDLTSAAAMYRGPVYALHDVSDKIPMTNSPILDPLPNLKIKVYNSSGAVT}
                    PQDDLAEFSSKLSPQMTQSLLENEALNLKNQSLARQTDPSCTAFGTFNSLGGHLIIPN
                    SGVSLLIPAGAIPOGRVYEMYVTVHRKENMRPPMEDSQTLLTPVVSCGPPGALLTRPV
                    ILTLHHCADPSTEDWKIQLKNQAVQGQWEDVVVVGEENFTTPCYIQLDAEACHILTEN
                    LSTYALVGQSTTKAAAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQM
                    GGQLLEEPKALRFKGSIHNLRLSIHDIAHSLWKSKLLAKYQEIPFYHIWSGSQRNLHC
                    TFTLERLSLNTVELVCKLCVROVEGEGOI FOLNCTVSEEPTGIDLPLLDPASTITTVT
                    GPSAFSIPLPIRQKLCSSLDAPQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDL
                    WEAQNFPDGNLSMLAAVLEEMGRHETVVYLAAEGQY"
ORIGIN
  Query Match
                         33.6%; Score 923.4; DB 3;
                                                      Length 3790;
 Best Local Similarity
                         61.3%; Pred. No. 3.9e-184;
 Matches 1659; Conservative
                                0; Mismatches 926;
                                                      Indels 120;
          149 TGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACA 208
Qу
                    275 TTCCATCTGACCCACCTGAGCCATTGCCACACTTCCTCATTGAGCCCGAGGAAGCTTACA 334
Db
          209 TCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCT 268
Qy
```

11 11 111

11 11:11 11:1:1: 11

1 11 11111111111

Db

| Qy | 269  | TCAAGTGCAACGGGGAGTGGGTGCCCAGGTGGACCACGTGATCGAGCGCAGCACAGACG  | 328  |
|----|------|--|------|
| Db | 395  | TCAAGTGCAACAGCGAGTGGGTTCATCAGAAGGACCACGTAGTAGACGAGAGAGTAGATG   | 454  |
| Qу | 329  | GGAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGA   | 388  |
| Db | 455  | AAACCTCTGGTCTAATTGTGAGAGAGAGTGAGCATTGAGATTTCACGCCAGCAGGTGGAGG  | 514  |
| Qу | 389  | AGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCA   | 448  |
| Db | 515  | AACTGTTTGGGCCTGAAGATTACTGGTGCCAGTGTGTGGCCTGGAGCTCAGCAGGCACTA   | 574  |
| Qу | 449  | CCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGC   | 508  |
| Db | 575  | CGAAGAGTCGGAAGGCATACGTGCGCATTGCGTATCTGCGGAAGACATTCGAGCAGGAAC   | 634  |
| Qу | 509  | CGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGG   | 568  |
| Db | 635  | CCTTGGGAAAGGAAGTGTCCTTGGAGCAGGAAGTCTTACTCCAGTGTCGGCCACCTGAAG   | 694  |
| Qy | 569  | GCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGG   | 628  |
| Db | 695  | GGATCCCAGTGGCTGAGGTAGAATGGCTAAAGAATGAAGACATAATTGATCCTGCTGAAG   | 754  |
| Qy | 629  | ACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTG   | 688  |
| Db | 755  | ATCGGAACTTTTATATTACTATCGATCACAACCTGATCATCAAGCAAG   | 814  |
| Qу | 689  | ACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCG   | 748  |
| Db | 815  | ATACAGCAAATTATACCTGTGTTGCCAAAAATATTGTTGCCAAGAGAAAAAAGCACCACAG  | 874  |
| Qу | 749  | CTGCTGTCATCGTCTACGTGAACGGTGGGTCGACGTGGACCGAGTGGTCCGTCTGCA  | 808  |
| Db | 875  | ${\tt CCACTGTCATCGTGTATGTTAATGGTGGCTGGTCCACCTGGACAGAGTGGTCTGTGTAATGGTGGTCTGTGTGTAATGGTGGTGGTCTGTGTGTAATGGTGGTGGTGTGTGT$ | 934  |
| Qy | 809  | GCGCCAGCTGTGGGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTC   | 868  |
| Db | 935  | ACAGCCGCTGTGGGCGAGGATATCAGAAACGCACAAGAACCTGCACCAACCCAGCCCCAC   | 994  |
| Qy | 869  | TCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGT   | 928  |
| Db | 995  | ${\tt TCAATGGTGGGGCCTTCTGTGAGGGGCAGAGTGTGCAGAAAATAGCATGCACTACGTTAT}$   | 1054 |
| Qy | 929  | GCCCAGTAGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCA   | 988  |
| Db | 1055 | GTCCAGTGGATGGTAGGTGGACTTCATGGAGCAAATGGTCAACCTGTGGGACTGAATGCA   | 1114 |
| Qy | 989  | CCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCC   | 1048 |
| Db | 1115 | CCCACTGGCGCAGGAGGGAGTGTACAGCACCAGCCCCCAAGAACGGGGGTAAGGACTGTG   | 1174 |
| Qy | 1049 | AGGGCACTGACCTGGACACCGCAACTGTACCAGTGACCTCTGTGTACACAG  | 1100 |
| Db | 1175 | ATGGCCTGGTCCTCCAATCCAAGAACTGCACTGATGGGCTGTGCATGCA  | 1234 |

| Qy | 1101 | TGCTTCTGGCC   | 1111 |
|----|------|---|------|
| Db | 1235 | ACCCCATTTCAACTGAGCACAGACCCCAGAATGAATATGGATTTTCTTCTGCTCCTGACT              | 1294 |
| Qу | 1112 | CTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGC                 | 1168 |
| Db | 1295 | CAGATGATGTGGCTCTCTACGTGGGGATTGTGATCGCTGTAACAGTCTGTCT                      | 1354 |
| Qу | 1169 | TGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGG              | 1228 |
| Db | 1355 | CTGTTGTGGTGGCCCTGTTTGTGTATCGGAAGAACCACCGTGACTTTGAGTCTGACATCA              | 1414 |
| Qу | 1229 | CTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAG              | 1288 |
| Db | 1415 | TTGACTCCTCAGCACTCAATGGCGGCTTTCAGCCTGTGAACATCAAGGCTG                       | 1465 |
| Qу | 1289 | ACAACCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCACCACCACCACCACCACC | 1348 |
| Db | 1466 | CCAGACAAGATCTCCTGGCTGTCCCCCCTGACCTCACCTCAGCTGCAGCCATGTACAGGG              | 1525 |
| Qy | 1349 | GCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGG                   | 1403 |
| Db | 1526 | GACCTGTCTATGCTCTGCATGATGTCTCAGACAAAATCCCAATGACCAACTCTCCAATTC              | 1585 |
| Qу | 1404 | GCACCTGCTCAGCCCCTGGGTGGCGGCCGCCACACTGCACCACAGCTCTCCCACCTC                 | 1463 |
| Db | 1586 | TGGACCCACTACCCAACTTGAAAATCAAAGTGTACAACAGCTCAGGTGCTGTCACTCCTC              | 1645 |
| Qу | 1464 | TGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCC                                       | 1498 |
| Db | 1646 | AGGATGACCTTGCCGAGTTCTCATCCAAACTGTCACCCCAGATGACCCAGTCCTTGCTAG              | 1705 |
| Qу | 1499 | AGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACA                                   | 1537 |
| Db | 1706 | AGAATGAGGCCCTTAACCTGAAGAACCAGAGCCTCGCAAGACAGAC                            | 1765 |
| Qy | 1538 | TGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCA              | 1597 |
| Db | 1766 | CAGCATTTGGTACCTTCAACTCTCTTGGGGGTCACCTCATCATTCCTAATTCAGGAGTAA              | 1825 |
| Qy | 1598 | GCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGC               | 1657 |
| Db | 1826 | GCTTGCTGATTCCCGCTGGGGCCATTCCTCAGGGGAGAGTCTATGAAATGTATGT                   | 1885 |
| Qy | 1658 | TGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTG                                  | 1717 |
| Db | 1886 | TACACAGGAAAGAAATATGAGGCCCCCCATGGAAGACTCTCAGACCCTACTTACCCCTG               | 1945 |
| QУ | 1718 | TCGTTAGCTGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACC               | 1777 |
| Db | 1946 | TGGTGAGCTGTGGGCCTCCTGGAGCTCTGCTGACCCGCCCTGTCATCCTCACTCTGCATC              | 2005 |
| Qy | 1778 | ACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGG              | 1837 |
| Db | 2006 |   | 2065 |
| Ov | 1838 | GCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCCTCCCACCTCTACTACTGCC             | 1897 |

| Db | 2066 |  | 2125 |
|----|------|--|------|
| Qу | 1898 | AGCTGGAGGCCAGTGCCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGG    | 1957 |
| Db | 2126 | AGCTGGATGCAGAGGCTTGCCATATCCTCACAGAGAACCTCAGTACCTATGCCCTGGTTG | 2185 |
| QУ | 1958 | GAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGG | 2017 |
| Db | 2186 |  | 2245 |
| Qу | 2018 | CCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC                   | 2077 |
| Db | 2246 | GCTGCTCTTCCCTGGAGTACAGCATTAGAGTCTACTGCCTGGATGACACACAGGATGCCC | 2305 |
| Qу | 2078 | TCAAGGAGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGG    | 2137 |
| Db | 2306 | TGAAGGAAGTTCTACAACTGGAGAGGCAAATGGGAGGACAGCTCCTAGAAGAACCCAAGG | 2365 |
| Qу | 2138 | TCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC                | 2197 |
| Db | 2366 | CTCTTCGTTTTAAAGGCAGCATCCACAACCTGCGCCTGTCTATTCATGACATCGCCCATT | 2425 |
| Qу | 2198 | CCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGA | 2257 |
| Db | 2426 | CCCTCTGGAAGAGCAAATTGCTGGCTAAGTATCAGGAAATTCCATTTTACCACATCTGGA | 2485 |
| Qy | 2258 | ATGGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTA  | 2317 |
| Db | 2486 | GTGGCTCTCAAAGAAACCTCCACTGCACCTTCACTCTGGAAAGACTCAGCCTAAACACAG | 2545 |
| Qу | 2318 | GTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCA | 2377 |
| Db | 2546 | TGGAACTGGTTTGCAAACTCTGTGTGCGGCAGGTTGAAGGAGAAGGGCAGATCTTCCAGC | 2605 |
| Qу | 2378 | TCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAG | 2437 |
| Db | 2606 | TCAACTGTACTGTCAGAGGAACCTACTGGCATCGACTTACCTCTCCTGGACCCTGCTA   | 2665 |
| Qу | 2438 | CGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGA | 2497 |
| Db | 2666 | GTACCATCACCACTGTCACCGGACCAAGTGCTTTCAGCATTCCTCTCCCTATCCGGCAGA | 2725 |
| Qу | 2498 | AGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCC | 2557 |
| Db | 2726 | AGCTATGCAGCAGCCTGGATGCCCCTCAAACAAGAGGCCATGACTGGAGGATGCTGGCCC | 2785 |
| Qy | 2558 | AGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCA  | 2617 |
| Db | 2786 | ATAAACTCAACCTGGACAGGTACTTGAATTACTTTGCCACCAAATCGAGCCCAACTGGCG | 2845 |
| Qу | 2618 | TGATCCTCAACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC        | 2677 |
| Db | 2846 | TAATCCTGGATCTTTGGGAAGCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGGCAG | 2905 |
| Qу | 2678 | CAGCAGTGGCTGGACTGGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGT      | 2737 |

```
Db
        2906 CCGTCCTGGAAGAAATGGGAAGACATGAGACAGTGGTGTACTTGGCAGCAGAAGGACAGT 2965
        2738 GCTGA 2742
Qy
               111
        2966 ATTGA 2970
Db
RESULT 4
CR598115
LOCUS
           CR598115
                                  1852 bp
                                            mRNA
                                                    linear
                                                             HTC 21-JUL-2004
           full-length cDNA clone CS0DA006YG16 of Neuroblastoma of Homo
DEFINITION
           sapiens (human).
ACCESSION
           CR598115
           CR598115.1 GI:50478922
VERSION
           HTC; CNSLT cDNA.
KEYWORDS
SOURCE
           Homo sapiens (human)
  ORGANISM
           Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
              (bases 1 to 1852)
 AUTHORS
           Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE
           Full-length cDNA libraries and normalization
  JOURNAL
           Unpublished
 REMARK
           Contact : Feng Liang Email : fliang@lifetech.com URL :
           http://fulllength.invitrogen.com/ InVitroGen Corporation 1600
           Faraday Avenue
REFERENCE
              (bases 1 to 1852)
 AUTHORS
           Genoscope.
  TITLE
           Direct Submission
  JOURNAL
           Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
           BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
           - Web : www.genoscope.cns.fr)
COMMENT
           1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
           end enriched, double-strand cDNA was digested with Not I and cloned
           into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
           was normalized. Library was constructed by Life Technologies, a
           division of Invitrogen.
FEATURES
                    Location/Qualifiers
                    1. .1852
    source
                    /organism="Homo sapiens"
                    /mol type="mRNA"
                    /db xref="taxon:9606"
                    /clone="CS0DA006YG16"
                    /tissue type="Neuroblastoma"
                    /plasmid="pCMVSPORT 6"
ORIGIN
 Query Match
                        32.1%; Score 884; DB 3; Length 1852;
 Best Local Similarity
                        100.0%; Pred. No. 7.5e-176;
 Matches 884; Conservative
                              0; Mismatches
                                                0; Indels
                                                              0; Gaps
                                                                         0;
Qу
        1869 GGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCAC 1928
             Db
           1 GGAGGCGCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCAC 60
Qy
        1929 CGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCG 1988
```

| Db | 61     | $\tt CGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCG$ | 120  |
|----|--------|--|------|
| Qу |        | CCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC                        |      |
| Db | 121    | CCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC                        | 180  |
| Qу | 2049   | CTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCT       | 2108 |
| Db | 181    | CTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCT       | 240  |
| Qу | 2109   | GGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCT        | 2168 |
| Db | . 241  | GGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCT       | 300  |
| Qу | 2169   | GCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTA       | 2228 |
| Db | 301    | GCGCCTATCCATCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTA        | 360  |
| QУ | 2229   | CCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTT       | 2288 |
| Db | 361    | CCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTT       | 420  |
| Qу | 2289   | CACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCA       | 2348 |
| Db | 421    | CACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCA       | 480  |
| Qу | 2349   | GGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGACACAAGGTT       | 2408 |
| Db | 481    | GGTGGAGGGCGACGGCAGAGCTTCAGCATCAACATCACCAAGGACACAAGGTT              | 540  |
| Qу | 2409   | TGCTGAGCTGCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGC             | 2468 |
| Db | 541    | TGCTGAGCTGCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGC             | 600  |
| Qу | 2469   | CTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCC            | 2528 |
| Db | 601    | CTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCC            | 660  |
| Qу | 2529   | GCGGGGTGCCGACTGGCCGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTT       | 2588 |
| Db | 661    | GCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTT       | 720  |
| Qу | 2589   | CTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGCACTT         | 2648 |
| Db | 721    | CTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGCACTT         | 780  |
| Qу | 2649   | CCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGCCA            | 2708 |
| Db | 781    |  | 840  |
| Qу | 2709   | TGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCC                         |      |
| Dh | . 9/11 | TGGCCTCTTCACAGTCTCGAGGCTGAGTGCTGAGGCCGGCC                          |      |

```
3866 bp
                                                         linear
                                                                   HTC 03-APR-2004
LOCUS
            AK018177
                                                 mRNA
            Mus musculus adult male medulla oblongata cDNA, RIKEN full-length
DEFINITION
            enriched library, clone:6330415E02 product:TRANSMEMBRANE RECEPTOR
            UNC5H2 homolog [Rattus norvegicus], full insert sequence.
ACCESSION
            AK018177
            AK018177.1 GI:12857775
VERSION
KEYWORDS
            HTC; CAP trapper.
            Mus musculus (house mouse)
SOURCE
  ORGANISM
            Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
  AUTHORS
            Carninci, P. and Hayashizaki, Y.
            High-efficiency full-length cDNA cloning
  TITLE
            Meth. Enzymol. 303, 19-44 (1999)
  JOURNAL
            99279253
  MEDLINE
            10349636
   PUBMED
REFERENCE
            2
            Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
  AUTHORS
            Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
  TITLE
            prepare full-length cDNA libraries for rapid discovery of new genes
            Genome Res. 10 (10), 1617-1630 (2000)
  JOURNAL
  MEDLINE
            20499374
   PUBMED
            11042159
REFERENCE
            Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
  AUTHORS
            Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
            Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
            Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
            Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
            Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
            Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
            RIKEN integrated sequence analysis (RISA) system--384-format
  TITLE
            sequencing pipeline with 384 multicapillary sequencer
            Genome Res. 10 (11), 1757-1771 (2000)
  JOURNAL
            20530913
  MEDLINE
            11076861
   PUBMED
REFERENCE
            The RİKEN Genome Exploration Research Group Phase II Team and the
  AUTHORS
            FANTOM Consortium.
            Functional annotation of a full-length mouse cDNA collection
  TITLE
  JOURNAL
            Nature 409, 685-690 (2001)
REFERENCE
            The FANTOM Consortium and the RIKEN Genome Exploration Research
  AUTHORS
            Group Phase I & II Team.
  TITLE
            Analysis of the mouse transcriptome based on functional annotation
            of 60,770 full-length cDNAs
  JOURNAL
            Nature 420, 563-573 (2002)
REFERENCE
                (bases 1 to 3866)
            Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H.,
  AUTHORS
            Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
            Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
            Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
            Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
            Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
```

Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,

```
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
      Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
      Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
      Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
      Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
      Direct Submission
      Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
      Physical and Chemical Research (RIKEN), Laboratory for Genome
      Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
      RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
      Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp,
      URL: http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
      Fax:81-45-503-9216)
      Please visit our web site (http://genome.gsc.riken.jp/) for further
      details.
      cDNA library was prepared and sequenced in Mouse Genome
      Encyclopedia Project of Genome Exploration Research Group in Riken
      Genomic Sciences Center and Genome Science Laboratory in RIKEN.
      Division of Experimental Animal Research in Riken contributed to
      prepare mouse tissues. First strand cDNA was primed with a primer
      prepared by using trehalose thermo-activated reverse transcriptase
      and subsequently enriched for full-length by cap-trapper. cDNA went
      through one round of normalization to Rot = 10.0 and subtraction to
      Rot = 100.0. Second strand cDNA was prepared with the primer
      adapter of sequence [5'
      with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
      bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
      end: BamHI. Host: DH10B.
               Location/Qualifiers
               1. .3866
source
               /organism="Mus musculus"
               /mol type="mRNA"
               /strain="C57BL/6J"
               /db xref="FANTOM DB:6330415E02"
               /db xref="taxon:10090"
               /clone="6330415E02"
               /sex="male"
               /tissue_type="medulla oblongata"
               /clone lib="RIKEN full-length enriched mouse cDNA library"
               /dev stage="adult"
               417. .3254
               /note="unnamed protein product; TRANSMEMBRANE RECEPTOR
               UNC5H2 homolog [Rattus norvegicus] (SPTR|008722, evidence:
               FASTY, 96.5%ID, 100%length, match=2835)
               putative"
               /codon start=1
               /protein id="BAB31108.1"
               /db xref="GI:12857776"
               translation="MRARSGVRSALLLALLLCWDPTPSLAGVDSAGQVLPDSYPSAPA/
               EQLPYFLLEPQDAYIVKNKPVELHCRAFPATQIYFKCNGEWVSQNDHVTQESLDEATG
               LRVREVQIEVSRQQVEELFGLEDYWCQCVAWSSSGTTKSRRAYIRIAYLRKNFDQEPL
               AKEVPLDHEVLLQCRPPEGVPVAEVEWLKNEDVIDPAQDTNFLLTIDHNLIIRQARLS
               DTANYTCVAKNIVAKRRSTAATVIVYVNGGWSSWAEWSPCSNRCGRGWQKRTRTCTNP
               APLNGGAFCEGOAFQKTACTTVCPVDGAWTEWSKWSACSTECAHWRSRECMAPPPQNG
```

GRDCSGTLLDSKNCTDGLCVLNQRTLNDPKSHPLETSGDVALYAGLVVAVFVVVAVLM

TITLE

COMMENT

**FEATURES** 

CDS

JOURNAL

AEGVIVYRRNCRDFDTDITDSSAALTGGFHPVNFKTARPNNPQLLHPSAPPDLTASAG IYRGPVYALQDSADKIPMTNSPLLDPLPSLKIKVYNSSTIGSGSGLADGADLLGVLPP GTYPGDFSRDTHFLHLRSASLGSQHLLGLPRDPSSSVSGTFGCLGGRLSLPGTGVSLL VPNGAIPQGKFYDLYLHINKAESTLPLSEGSQTVLSPSVTCGPTGLLLCRPVVLTVPH CAEVIAGDWIFQLKTQAHQGHWEEVVTLDEETLNTPCYCQLEAKSCHILLDQLGSYVF MGESYSRSAVKRLQLAIFAPALCTSLEYSLRVYCLEDTPVALKEVLELERTLGGYLVE EPKPLLFKDSYHNLRLSLHDIPHAHWRSKLLAKYQEIPFYHVWNGSQRALHCTFTLER HSLASTEFTCKVCVRQVEGEGQIFQLHTTLAETPAGSLDALCSAPGNAITTQLGPYAF KIPLSIRQKICSSLDAPDSRGNDWRLLAQKLSMDRYLNYFATKASPTGVILDLWEARQ QDDGDLNSLASALEEMGKSEMLVAMATDGDC"

## ORIGIN

|         | cal : | 31.7%; Score 872.4; DB 3; Length 3866; Similarity 60.6%; Pred. No. 2.3e-173; 6; Conservative 0; Mismatches 941; Indels 129; Gaps | 8;   |
|---------|-------|--|------|
| Macches | 104   | o, conservative o, mismatches 941, indeis 129, Gaps  | 0,   |
| Qу      | 157   | GCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAG   | 216  |
| Db      | 540   | GCCCTGCGGAGCAGCTGCCGTACTTCCTATTGGAGCCACAGGACGCCTACATCGTAAAG  | 599  |
| Qy      | 217   | AACAAGCCAGTGCTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCŢTCAAGTGC   | 276  |
| Db      | 600   | ${\tt AACAAGCCAGTGGAACTGCAGTGCAGAGCCTTCCCTGCCACGCAGATCTACTTCAAGTGT}$   | 659  |
| Qу      | 277   | AACGGGGAGTGGGTGCCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGT  | 336  |
| Db      | 660   | AATGGCGAGTGGGTCAGCCAGAATGACCACGTCACACAGGAGAGCCTGGATGAGGCCACA   | 719  |
| Qy      | 337   | GGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTC   | 396  |
| Db      | 720   | GGCTTGCGGGTGCGAGAGTGCAGATCGAGGTGTCACGCAGCAAGTGGAGGAACTCTTC   | 779  |
| Qу      | 397   | GGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGT   | 456  |
| Db      | 780   | GGGCTCGAGGACTACTGGTGCCAGTGCGTGGCCTGGAGCTCTTCGGGAACTACCAAGAGT   | 839  |
| Qу      | 457   | CAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCC   | 516  |
| Db      | 840   | CGCCGAGCCTACATCCGCATTGCCTACTTGCGCAAGAACTTTGACCAGGAGCCTCTGGCC   | 899  |
| QУ      | 517   | AAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCT   | 576  |
| Db      | 900   | AAGGAGGTACCCTTGGATCATGAGGTCCTTCTGCAGTGCCGCCCACCGGAGGGAG  | 959  |
| Qу      | 577   | CCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAAT   | 636  |
| Db      | 960   | GTGGCTGAGGTGGAATGGCTCAAGAATGAAGATGTCATTGACCCCGCTCAGGACACTAAC   | 1019 |
| Qу      | 637   | GTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCC   | 696  |
| Db      | 1020  | TTCCTGCTCACCATTGACCACAACCTCATCATCCGCCAGGCGCGCCTCTCAGACACGGCC   | 1079 |
| Qу      | 697   | AACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTC   | 756  |
| Db      | 1080  | AACTACACCTGTGTGGCCAAGAATATCGTGGCCAAGCGCCGGAGCACCGCGGCCACAGTC   | 1139 |
| Qу      | 757   | ${\tt ATCGTCTACGTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGC}$   | 816  |

| Db     | 1140 | ATCGTCTATGTGAATGGAGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGTTCCAATCGC | 1199 |
|--------|------|--|------|
| Qу     | 817  | TGTGGGCGCGGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGG  | 876  |
| Db     | 1200 | TGTGGCCGAGGCTGCAGAAGCGTACTCGGACCTGCACCAATCCAGCCCCACTCAATGGA  | 1259 |
| Qу     | 877  | GGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTA | 936  |
| Db     | 1260 | GGCGCCTTCTGTGAGGGACAGGCCTTCCAGAAGACAGCTTGCACCACCGTGTGCCCAGTG | 1319 |
| QУ     | 937  | GACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGG | 996  |
| Db     | 1320 | GATGGAGCGTGGACCGAGTGGAGCAAGTGGTCTGCCTGCAGCACAGAGTGTGCGCACTGG | 1379 |
| Qу     | 997  | CGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACT | 1056 |
| Db     | 1380 | CGCAGCCGCGAGTGCATGGCACCGCCACCCCAGAACGGAGGCCGTGACTGCAGCGGGACG | 1439 |
| QУ     | 1057 | GACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCC      | 1111 |
| Db     |      | CTACTTGACTCCAAGAACTGCACTGATGGGCTGTGCGTGC                     |      |
| Qу     | 1112 | CTGAGGACGTGGCCCTCTATGTGGGCCTCATC                             | 1143 |
| Db.    | 1500 | GACCCTAAAAGCCACCCCTGGAGACATCGGGAGATGTGGCACTGTACGCAGGCCTTGTG  | 1559 |
| Qy.    | 1144 | GCCGTGGCCGTCTGCTGCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGG        | 1200 |
| Db     |      | GTGGCCGTCTTTGTGGTGGTAGCGGTTCTCATGGCCGAGGGAGTGATCGTATACCGGAGA |      |
| QУ     |      | AAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC     |      |
| Db     |      | AACTGCCGGGACTTCGACACGGACATCACCGACTCCTCTGCGGCCCTCACTGGTGGCTTC |      |
| QУ     |      | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATC       |      |
| Db     |      | CACCCTGTCAACTTCAAGACTGCAAGGCCCAACAACCCGCAGCTCCTGCACCCGTCCGCC |      |
| QУ     |      | CAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGA-    |      |
| Db     |      | CCTCCAGACCTAACGGCCAGTGCTGGCATCTACCGCGGGCCTGTGTATGCCCTGCAGGAC |      |
| Qy<br> | 1371 |  |      |
| Db     |      | TCCGCCGACAAGATCCCCATGACTAATTCGCCCCTGCTGGATCCCCTGCCCAGCCTCAAG |      |
| Qу     |      | TTCCAGCTCACCAATGGGCACCTGCTCAGC                               |      |
| Ďр     |      | ATCAAGGTCTATAACTCCAGCACCATCGGTTCTGGGTCTGGCCTGATGGAGCCGAC     |      |
| QУ     |      | CCCCTGGGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAG    |      |
| Db     |      | CTGCTGGGTGTCCTCCCGCCGGGCACGTACCCAGGCGATTTCTCCCGGGACACCCATTTC |      |
| Qу     | 1477 | TTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACC       | 1530 |
|        |      |  |      |

| Db | 1980 | CTGCACCTGCGCAGTGCCAGCCTTGGTTCCCAGCACCTCCTGGGCCTACCTCGGGACCCC  | 2039 |
|----|------|---|------|
| Qy | 1531 | AGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGGCCGGCTGATGATCCCTAATACA | 1590 |
| Db | 2040 | AGCAGCAGTGTCAGCGGCACCTTTGGTTGCCTGGGAGGAAGGCTGAGCCTCCCCGGCACA  | 2099 |
| Qу | 1591 | GGTATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTAC   | 1650 |
| Db | 2100 | GGGGTCAGCCTGTTGGTACCAAATGGAGCCATTCCCCAGGGCAAGTTCTATGACCTGTAT  | 2159 |
| Qу | 1651 | CTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTG               | 1710 |
| Db | 2160 | CTACATATCAACAAGGCCGAAAGCACCCTCCCACTTTCAGAAGGTTCCCAGACAGTATTG  | 2219 |
| QУ | 1711 | AGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCT  | 1770 |
| Db | 2220 | AGCCCTCGGTGACCTGTGGGCCCACAGGCCTACTCCTGTGCCGCCCTGTCGTCCTCACC   | 2279 |
| Qу | 1771 | ATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCG  | 1830 |
| Db | 2280 | GTGCCCCACTGTGCTGAAGTCATCGCTGGAGACTGGATCTTTCAGCTCAAGACCCAGGCC  | 2339 |
| QУ | 1831 | TGCGAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTAC  | 1890 |
| Db | 2340 | CATCAGGGCCACTGGGAGGAGGTGGTGACCTTGGATGAGGAGACCCTCAACACCCCTGC   | 2399 |
| Qу | 1891 | TACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCC  | 1950 |
| Db | 2400 | TACTGCCAGCTGGAGGCTAAGTCCTGCCACATCCTGCTGGACCAGCTGGGTTCCTACGTA  | 2459 |
| Qу | 1951 | CTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCG  | 2010 |
| Db | 2460 | TTCATGGGCGAGTCCTACTCTCGCTCTGCAGTCAAGCGGCTCCAGCTGGCCATCTTCGCC  | 2519 |
| Qу | 2011 | CCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC             | 2070 |
| Db | 2520 | CCAGCCCTCTGCACCTCCCTGGAGTATAGCCTCAGGGTCTACTGTCTGGAGGACACACCT  | 2579 |
| Qу |      | GATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAG | 2130 |
| Db |      | GTAGCACTGAAGGAGGTCCTGGAGCTGGAGAGGACTCTGGGTGGCTACTTGGTGGAGGAG  | 2639 |
| Qу | 2131 | CCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC          | 2190 |
| Db | 2640 | CCCAAGCCTTTGCTCTTTAAGGACAGTTACCACAACCTACGCCTCTCCCTCC          | 2699 |
| Qу | 2191 | CCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCAC  | 2250 |
| Db | 2700 | CCCCATGCCCACTGGAGGAGCAAACTACTGGCCAAGTACCAGGAGATTCCCTTCTACCAC  | 2759 |
| Qу | 2251 | ATCTGGAATGGCACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCC  | 2310 |
| Db | 2760 | GTCTGGAATGGCAGCCAGAGAGCCCTGCACTGCACTTTCACCCTGGAGAGGCATAGCCTG  | 2819 |
| Qy | 2311 | AGCACTAGTGACCTGCCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGCAGAGC       | 2370 |
| Db | 2820 | GCCTCCACGGAGTTCACCTGTAAGGTCTGCGTGCGGCAGGTCGAAGGGGAAGGCCAGATT  | 2879 |

```
2371 TTCAGCATCAACTTCAACATCACC---AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTG 2427
Qy
                                       1 1
                                              1 1 1
                                                              111
                    1 11
                           | | |
Db
        2880 TTCCAGCTGCACACACGTTGGCCGAGACGCCTGCTGGCTCCCTGGATGCTCTCTGCTCT 2939
        2428 GAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTC 2487
Qу
                             2940 GCCCGGGCAATGCCATCACCAGCTGGGACCCTATGCCTTCAAGATACCCCTGTCC 2999
Db
        2488 ATTCGGCAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGG 2547
Qу
             11 11 11 1111
                                1111111111 1 111
                                                      3000 ATCCGCCAAAAGATCTGCAGCAGCCTGGACGCCCCGACTCCCGGGGCAACGACTGGAGG 3059
Db
        2548 ACTCTGGCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGC 2607
Qу
                                1111 11111 11
Db
        3060 CTGTTGGCGCAGAAGCTGTCCATGGACCGGTACCTAAACTACTTCGCCACCAAAGCTAGT 3119
        2608 CCCACAGCCATGATCCTCAACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGC 2667
Qy
                      3120 CCCACAGGTGTCATCTTAGACCTCTGGGAAGCTCGGCAACAGGATGACGGGGACCTCAAC 3179
Db
        2668 CAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAGACGCTGGCCTCTTCACAGTGTCG 2727
Qу
                         1111111
                                               -11
        3180 AGCCTGGCCAGTGCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGGTAGCCATGGCCACA 3239
Db
        2728 GAGGCTGAGTGCTGAG 2743
Qу
             11 : 11 :11111
Db
        3240 GATGGCGATTGCTGAG 3255
RESULT 6
AY406492
                                  2802 bp
                                                             GSS 15-DEC-2003
LOCUS
           AY406492
                                             DNA
                                                     linear
           Pan troglodytes UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION
           genomic survey sequence.
ACCESSION
           AY406492
           AY406492.1 GI:39762466
VERSION
KEYWORDS
           GSS.
           Pan troglodytes (chimpanzee)
SOURCE
           Pan troglodytes
 ORGANISM
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
              (bases 1 to 2802)
           Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
  AUTHORS
           Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
           Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
           Adams, M.D. and Cargill, M.
  TITLE
           Inferring nonneutral evolution from human-chimp-mouse orthologous
  JOURNAL
           Science 302 (5652), 1960-1963 (2003)
   PUBMED
           14671302
             (bases 1 to 2802)
REFERENCE
  AUTHORS
           Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
           Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
           Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
           Adams, M.D. and Cargill, M.
  TITLE
           Direct Submission
```

```
Rockville, MD 20850, USA
         This sequence was made by sequencing genomic exons and ordering
COMMENT
         them based on alignment.
               Location/Oualifiers
FEATURES
                1. .2802
   source
                /organism="Pan troglodytes"
                /mol_type="genomic DNA"
                /db xref="taxon:9598"
                <1. .>2802
   gene
                /gene="UNC5C"
                /locus tag="HCM2575"
ORIGIN
                   29.6%; Score 814; DB 9; Length 2802;
 Query Match
                   53.8%; Pred. No. 4.7e-161;
 Best Local Similarity
 Matches 1477; Conservative
                         0; Mismatches 1200; Indels
                                                69; Gaps
                                                          6;
        57 GCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGG 116
Qy
                         1
        66 GCTCGTGCTGCCCTGCCCTGCTCAGCGCCAGCGCACTGGCTCCGCCGCCCAAGA 125 ·
Db
       117 TGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCC 176
Qy
                                             11 11 1 11 11
                   - 1
                         1
                           1
       126 TGATGACTTTTTCATGAACTCCCAGAAACTTTTCCTTCGGATCCACCTGAGCCTCTGCC 185
Db
       177 CCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGT 236
Qy
           186 ACATTTCCTTATTGAACCTGAGGAAGCTTATATTGTGAAGAATAAGCCTGTGAACCTGTA 245
Db
       237 GTGCAAGGCC-----GTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGT 290
Qy
       Db
       291 GCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCAT 350.
Qy
       Db
       351 GGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATA 410
Qу
       Db
       411 CTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACAT 470
Qy
       Db
       471 CCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCT 530
Qу
                       Db
       531 GGAGCAGGGCATCGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGA 590
Qу
                   +1
       546 GGAACAGGAAGTCTTACTCCAGTGTCGGCCACCTGAAGGGATCCCAGTGGCTGAGNNNNN 605
Db
       591 GTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCG 650
Qy
                  11 11 111 1 1 11 11
                                      11 1 111 1 11 11 11
       606 NNNNNGAAAAATGAAGACATAATTGATCCCGTTGAAGATCGGAATTTTTATATTACTAT 665
Db
```

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,

**JOURNAL** 

|   | Qy              | 651  | GGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGT | 710  |
|---|-----------------|------|--|------|
|   | Db              | 666  | TGATCACAACCTCATCATAAAGCAGGCCCGACTCTCTGATACTGCAAATTATACCTGTGT | 725  |
|   | Qy              | 711  | GGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAA | 770  |
|   | Db              | 726  | TGCCAAAAACATTGTTGCCAAGAGGAAAAGTACAACTGCCACTGTCATAGTCTATGTCAA | 785  |
|   | Qy              | 771  | CGGTGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTG  | 830  |
|   | Db              | 78.6 | CGGTGGCTGGTCCACCTGGACGGAGTGGTCTGTGTGTAACAGCCGCTGTGGACGAGGGTA | 845  |
|   | Qy              | 831  | GCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGA | 890  |
|   | Db              | 846  | TCAGAAACGTACAAGGACTTGTACCAACCCGGCACCACTCAATGGGGGTGCCTTCTGTGA | 905  |
|   | Qy              | 891  | GGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAG  | 950  |
|   | Db              | 906  | AGGGCAGAGTGTGCAGAAAATAGCCTGTACTACGTTATGCCCAGTGGATGGCAGGTGGAC | 965  |
|   | Qу              | 951  | CCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGACTGCACCCACTGGCGGAGCCGTGAGTG  | 1010 |
|   | Db              | 966  | GCCATGGAGCAAGTGGTCTACTTGTGGAACTGAGTGCACCCACTGGCGCAGGAGGGAG   | 1025 |
|   | Qy <sub>.</sub> | 1011 | CTCTGACCCAGCACCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCG  | 1070 |
|   | Db              | 1026 | CACGGCGCCAGCCCCCAAGAATGGAGGCAAGGACTGCGATGGCCTCGTCTTGCAATCCAA | 1085 |
|   | Qу              | 1071 | CAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTA | 1130 |
|   | Db .            | 1086 | GAACTGCACTGATGGGCTTTGCATGCAGACTGCTCCTGATTCAGATGATGTTGCTCTCTA | 1145 |
|   | Qy              | 1131 | TGTGGGCCTCATCGCCGTGGCCGTCTGCCTGCTGCTGCTGCTGCTCATCCT          | 1187 |
|   | Db              | 1146 | TGTTGGGATTGTGATAGCAGTGATCGTTTGCCTGGCGATCTCTGTAGTTGTGGCCTTGTT | 1205 |
|   | Qy              | 1188 | CGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCAC | 1247 |
|   | Db              | 1206 | TGTGTATCGGÄÄĞÄATCATCGTGACTTTGAGTCAGATATTATTGACTCTTCGGCACTCAA | 1265 |
|   | Qу              | 1248 | CTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCAC | 1307 |
|   | Db              | 1266 | TGGGGGCTTTCAGCCTGTGAACATCAAGGCAGCAAGACAAGA                   | 1316 |
| • | Qу              | 1308 | CATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCCCGGCA     | 1367 |
|   | Db              | 1317 | TGTACCCCCAGACCTCACGTCAGCTGCAGCCATGTACAGAGGACCTGTCTATGCCCTGCA | 1376 |
|   | Qy              | 1368 | GGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCC          | 1418 |
|   | Db              | 1377 | TGACGTCTCAGACAAAATCCCAATGACCAACTCTCCAATTCTGGATCCACTGCCCAACCT | 1436 |
|   | Qy              | 1419 | CCTGGGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTT    | 1478 |
|   | Db              | 1437 | GAAAATCAAAGTGTACAACACCTCAGGTGCTGTCACCCCCCAAGATGACCTCTCTGAGTT | 1496 |

|   | Qy   | 14/9 | CGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACAT  | 1538 |
|---|------|------|---|------|
|   | Db   | 1497 | TACGTCCAAGCTGTCCCCTCAGATGACCCAGTCATTGTTGGAGAATGAAGCCCTCAGCCT  | 1556 |
|   | Qy   | 1539 | GACCTATGGGACCTTCAA  | 1556 |
| · | Db   | 1557 | GAAGAACCAGAGTCTAGCAAGGCAGACTGATCCATCCTGTACCGCATTTGGCAGCTTCNN  | 1616 |
|   | Qу   | 1557 | CTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGA   | 1616 |
| • | Db   | 1617 |   | 1676 |
|   | Qy   | 1617 | TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGT  | 1676 |
|   | Db   | 1677 |   | 1736 |
|   | Qy   | 1677 | GAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC  | 1736 |
|   | Db   | 1737 | GAGGCCACCCATGGATGACTCTCAGACACTTTTGACCCCTGTGGTGAGCTGTGGGCCCCC  | 1796 |
|   | Qу   | 1737 | TGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCC  | 1796 |
|   | Db   | 1797 | AGGAGCTCTGCTCACCCGCCCCGTCGTCCTTACTATGCATCACTGCGCAGACCCCAATAC  | 1856 |
| • | Qy   | 1797 | TGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCT  | 1856 |
|   | Db   | 1857 |   | 1916 |
|   | Qy   | 1857 | GCACCTGGGCGAGGAGGCCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTG   | 1916 |
|   | Db   | 1917 | GGTGGTCGGGGAGAAACTTCACCACCCCCTGCTACATTCAGCTGGATGCAGAGGCCTG    | 1976 |
|   | Qy . | 1917 | CTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGC  | 1976 |
|   | Db   | 1977 | CCACATCCTCACAGAGAACCTCAGCACCTACGCCCTGGTAGGACATTCCACCACCAAAGC  | 2036 |
|   | Qу   | 1977 | TGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC       | 2036 |
|   | Db   | 2037 | GGCTGCGAAGCGCCTCAAGCTGGCCATCTTTGGGCCCCTGTGCTGCTCCTCGCTGGAGTA  | 2096 |
|   | Qy   | 2037 | CAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCT  | 2096 |
|   | Db   | 2097 | CAGCATCCGAGTCTACTGTCTGGATGACACCCAGGATGCCCTGAAGGAAATTTTACATCT  | 2156 |
|   | Qу   | 2097 | GGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAG  | 2156 |
|   | Db   | 2157 | TGAGAGACAGATGGGAGGACAGCTCCTAGAAGAACCTAAGGCTCTTCATTTTAAAGGCAG  | 2216 |
|   | Qy   | 2157 | TTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCT  | 2216 |
|   | Db   | 2217 | CACCCACAACCTGCGCCTGTCAATTCACGATATCGCCCCATTCCCTCTGGAAGAGCAAATT | 2276 |
|   | Qy   | 2217 | CCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTT  | 2276 |
|   | Db   | 2277 |   | 2336 |
|   | Ov   | 2277 | GCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCT  | 2336 |

```
Db
       2337 GCACTGCACCTTCACTCTGGAAAGATTTAGCCTGAACACAGTGGAGCTGGTTTGCAAACT 2396
       2337 GTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAA 2396
Qу
            2397 CTGTGTGCGGCAGGTGGAAGGAGAAGGGCAGATCTTCCAGCTCAACTGCACCGTGTCAGA 2456
Db
       2397 GGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGT 2456
Qу
           11 111 1 11111
                                       - 11
                                                  11 1 1 1
       2457 GGAACCTACTGGCATCGATTTGCCGCTGCTGGATCCTGCGAACACCATCACCACGGTCAC 2516
Db
       2457 GGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGA 2516
Qу
           2517 GGGGCCCAGTGCTTTCAGCATCCCTCTCCCTATCCGGCAGAAGCTCTGTAGCAGCCTGGA 2576
Db
       2517 CCCACCCTGTAGGCGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAG 2576
Qy
                   Db
       2577 CCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGGGA 2636
Qу
               2637 GTACTTGAATTACTTTGCCACCAAATCCAGCCCAACTGGCGTAATCCTGGATCTTTGGGA 2696
Db
       Qу
            Db
       2697 AGCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGGCAGCTGTCTTGGAAGAAATGGG 2756
       2697 CCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2742
Qу
              +++
       2757 AAGACATGAAACGGTGGTGTCCTTAGCAGCAGAAGGGCAGTATTAA 2802
Db
RESULT 7
AY411747
LOCUS
         AY411747
                             2532 bp
                                                    GSS 12-DEC-2003
                                      DNA
                                             linear
DEFINITION
         Homo sapiens HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
         genomic survey sequence.
ACCESSION
         AY411747
         AY411747.1 GI:39767715
VERSION
KEYWORDS
         GSS.
SOURCE
         Homo sapiens (human)
 ORGANISM
         Homo sapiens
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
            (bases 1 to 2532)
 AUTHORS
         Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
         Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
         Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
         Adams, M.D. and Cargill, M.
 TITLE
         Inferring nonneutral evolution from human-chimp-mouse orthologous
         gene trios
         Science 302 (5652), 1960-1963 (2003)
 JOURNAL
  PUBMED
         14671302
         .2 (bases 1 to 2532)
REFERENCE
 AUTHORS
         Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
         Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
         Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
```

```
Direct Submission
 TITLE
 JOURNAL
          Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
          Rockville, MD 20850, USA
          This sequence was made by sequencing genomic exons and ordering
COMMENT
          them based on alignment.
FEATURES
                 Location/Qualifiers
                  1. .2532
    source
                  /organism="Homo sapiens"
                  /mol type="genomic DNA"
                  /db xref="taxon:9606"
                  <1. .>2532
    gene
                 /locus tag="HCM4327"
ORIGIN
 Query Match
                      29.4%;
                             Score 810.4; DB 9;
                                              Length 2532;
                      61.1%; Pred. No. 2.7e-160;
 Best Local Similarity
 Matches 1540; Conservative
                            0; Mismatches 851;
                                                     129;
                                                                  9;
                                              Indels
                                                           Gaps
        352 GAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATAC 411
Qy
            13 GAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTCTTTGGGCTGGAGGATTAC 72
Db
        412 TGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCAAGAGTCAGAAGGCCTACATC 471
Qу
           73 TGGTGCCAGTGCCTGGAGCTCCGCGGGCACCACAAGAGTCGCCGAGCCTACGTC 132
Db
        472 CGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTG 531
Qу
                       133 CGCATCGCCTACCTGCGCAAGAACTTCGATCAGGAGCCTCTGGGCAAGGAGGTGCCCCTG 192
Db
        532 GAGCAGGGCATCGTGCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAG 591
Qy
                      193 GACCATGAGGTTCTCCTGCAGTGCCGCCGGAGGGGGGTGCCTGTGGCCGAGGTGGAA 252
Db
        592 TGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGG 651
Qy
            253 TGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACCAACTTCCTGCTCACCATC 312
Db
        652 GAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTG 711
Qy
                      313 GACCACAACCTCATCATCCGCCAGGCCCGCCTGTCGGACACTGCCAACTATACCTGCGTG 372
Db
        712 GCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAAC 771
Qу
            11111111111111
                             Db
        373 GCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACCGTCATCGTCTACGTGAAT 432
        772 GGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGG 831
Qу
                        11 11 11111 1
        433 GGCGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGCTCCAACCGCTGTGGCCGAGGCTGG 492
Db
        832 CAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAG 891
Qy
            11111 | 1 | 1 | 1111 | 11111 | 1111 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11
        493 CAGAAGCGCACCCGGACCTGCACCAACCCCGCTCCACTCAACGGAGGGGCCTTCTGCGAG 552
Db
```

892 GGGCAGAATGTCCAGAAAACAGCCTGCGCCACCTGTGCCCAGTAGACGGCAGCTGGAGC 951 

Adams, M.D. and Cargill, M.

Qv

11 111

| Db   | 553  | GGCCAGGCATTCCAGAAGACCGCCTGCACCACCATCTGCCCAGTCGATGGGGCGTGGACG | 612  |
|------|------|--|------|
| Qу   | 952  | CCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGC | 1011 |
| Db   | 613  | GAGTGGAGCAAGTGGTCAGCCTGCAGCACTGAGTGTGCCCACTGGCGTAGCCGCGAGTGC | 672  |
| Qу   | 1012 | TCTGACCCAGCACCCGCAACGGAGGGGGGGGGGGGGGGG                      | 1071 |
| Db   | 673  | ATGGCGCCCCACCCCAGAACGGAGGCCGTGACTGCAGCGGGACGCTGCTCGACTCTAAG  | 732  |
| Qу   | 1072 | AACTGTACCAGTGACCTCTGTGTACACAGTGCTTCT                         | 1107 |
| Db   | 733  | AACTGCACAGATGGGCTGTGCATGCAAAATAAGAAAACTCTAAGCGACCCCAACAGCCAC | 792  |
| QУ   | 1108 | GGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTC             | 1155 |
| Db · | 793  | CTGCTGGAGGCCTCAGGGGATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCATCTTCGTG | 852  |
| QУ   | 1156 | TGCCTGGTCCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTG    | 1215 |
| Db   | 853  | GTCGTGGCAATCCTCATGGCGGTGGGGGTGGTGTACCGCCGCAACTGCCGTGACTTC    | 912  |
| QУ   | 1216 | GACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATC    | 1272 |
| Db   | 913  | GACACAGACATCACTGACTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTTT      | 972  |
| Qy   |      | AAGCCCAGCAAAGCAGACCCCCATCTGCTCACCATCCAGCCGGACCTCAGC          |      |
| Db   |      | AAGACGGCAAGCCCAGCACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTGACA   |      |
| Qу   |      | ACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGG                |      |
| Db   |      | GCCAGCGCCGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCCACCGACAAAATC  |      |
| Qу   |      | CCCAGCCCCAAGTTCCAGCTCACCA                                    |      |
| Db   |      | CCCATGACCAACTCTCCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGC   |      |
| Qу   |      | ATGGGCACCTGCTCAGCCCCCTGGGTGGCGCCCCCACACA                     |      |
| Db · |      | TCCAGCACCACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGGGTCTTG |      |
| Qу   |      | CTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG  |      |
| Db   |      | CCGCCTGGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGCGCAGC |      |
| Qy   |      | AACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT                |      |
| Db   |      | GCCAGCCTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGGGAGCAGCGTCAGC |      |
| ДУ   |      | GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC |      |
| Db   | 4    | GGCACCTTTGGCTGCCTGGGTGGGAGGCTCAGCATCCCCGGCACAGGGGTCAGCTTGCTG |      |
| Qу   |      | ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG  |      |
| Db   | T393 | GTGCCCAATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAG | 1452 |

| Qу |      | CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC  |       |
|----|------|---|-------|
| Db | 1453 | GCAGAAAGTACCCTCCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCGGTGACC  | 1512  |
| Qу | 1726 | TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG   | 1785  |
| Db | 1513 | TGTGGACCCACAGGCCTCCTGCTGTGCCGCCCCGTCATCCTCACCATGCCCCACTGTGCC  | 1572  |
| Qу | 1786 | GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG  | 1845  |
| Db | 1573 | GAAGTCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTGG  | 1632  |
| Qy | 1846 | GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG  | 1905  |
| Db | 1633 | GAGGAGGTGGTGACCCTGGATGAGGAGACCCTGAACACCCTGCTACTGCCAGCTGGAG    | 1692  |
| Qу | 1906 | GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC  | 1965  |
| Db | 1693 | CCCAGGGCCTGTCACATCCTGCTGGACCAGCTGGGCACCTACGTGTTCACGGGCGAGTCC  | 1752  |
| Qу | 1966 | CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC  | 2025  |
| Db | 1753 | TATTCCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCCGTCTTCGCCCCCGCCCTCTGCACC  | 1812  |
| Qу | 2026 | TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG  | 2085  |
| Db | 1813 | TCCCTGGAGTACAGCCTCCGGGTCTACTGCCTGGAGGACACGCCTGTAGCACTGAAGGAG  | 1872  |
| Qу | 2086 | GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC | 2145  |
| Db | 1873 | GTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATG  | 1932  |
| Qy | 2146 | TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG  | 2205  |
| Db | 1933 | TTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGG  | 1992  |
| Qу | 2206 | AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG  | 2,265 |
| Db | 1993 | AGGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTATCACATTTGGAGTGGCAGC  | 2052  |
| Qу | 2266 | CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG  | 2325  |
| Db | 2053 | CAGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTC  | 2112  |
| Qy | 2326 | GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC  | 2385  |
| Db | 2113 | ACCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACC  | 2172  |
| Qy | 2386 | AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGG     | 2442  |
| Db | 2173 |   | 2232  |
| Qу | 2443 | GTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATA  | 2502  |
| Db | 2233 |   | 2292  |

```
2503 ATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAA 2562
Qу
                  11111.11 1 111
                                      \Pi\Pi\Pi\Pi
                                               2293 TGCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAG 2352
Db
        2563 CTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATC 2622
Qу
                   2353 CTCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATC 2412
Db
        Qу
                                        11111
                2413 CTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGTGCC 2472
Db
        2683 GTGGCTGGACTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2742
Qу
                                                 - 11
                                          1
        2473 TTGGAGGAGATGGCAAGAGTGAGATGCTGGTGGCTGTGGCCACCGACGGGGACTGCTGA 2532
Db
RESULT 8
AY411749
                                                            GSS 12-DEC-2003
LOCUS
                                 2532 bp
                                            DNA
                                                   linear
           AY411749
          Mus musculus HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION
           genomic survey sequence.
ACCESSION
           AY411749
           AY411749.1 GI:39767717
VERSION
KEYWORDS
           GSS.
SOURCE
           Mus musculus (house mouse)
 ORGANISM Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
           1 (bases 1 to 2532)
           Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 AUTHORS
           Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
           Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
           Adams, M.D. and Cargill, M.
           Inferring nonneutral evolution from human-chimp-mouse orthologous
 TITLE
           gene trios
           Science 302 (5652), 1960-1963 (2003)
 JOURNAL
  PUBMED
           14671302
REFERENCE
             (bases 1 to 2532)
           Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 AUTHORS
           Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
           Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
           Adams, M.D. and Cargill, M.
           Direct Submission
 TITLE
           Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 JOURNAL
           Rockville, MD 20850, USA
COMMENT
           This sequence was made by sequencing genomic exons and ordering
           them based on alignment.
FEATURES
                   Location/Qualifiers
                   1. .2532
    source
                   /organism="Mus musculus"
                   /mol type="genomic DNA"
                   /db xref="taxon:10090"
                   <1. .>2532
    gene
                   /locus tag="HCM4327"
ORIGIN
```

|   |      | cal S | 28.4%; Score 780.4; DB 9; Length 2532; Similarity 60.1%; Pred. No. 5.8e-154; S; Conservative 0; Mismatches 876; Indels 129; Gaps | 8;   |
|---|------|-------|--|------|
|   | Qy   |       | GAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGAATAC  |      |
|   | Db   | 13    |  | 72   |
|   | Qу   | 412   | TGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATC   | 471  |
|   | Db   | 73    | TGGTGCCAGTGCGTGGCCTGGAGCTCTTCGGGAACTACCAAGAGTCGCCGAGCCTACATC   | 132  |
|   | Qу   | 472   | CGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTG   | 531  |
|   | Db   | 133   |  | 192  |
|   | Qy   | 532   | GAGCAGGGCATCGTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAG   | 591  |
|   | Db   | 193   |  | 252  |
|   | Qy   | 592   | TGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGG   | 651  |
|   | Db   | 253   | TGGCTCAAGAATGAAGATGTCATTGACCCCGCTCAGGACACTAACTTCCTGCTCACCATT   | 312  |
|   | Qу   | 652   | GAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTG   | 711  |
|   | Db   | 313   | GACCACAACCTCATCCGCCAGGCGCCCTCTCAGACACGGCCAACTACACCTGTGTG   | 372  |
|   | Qy   | 712   | GCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAAC   | 771  |
|   | Db   | 373   | GCCAAGAATATCGTGGCCAAGCGCCGGAGCACCACGGCCACAGTCATCGTCTATGTGAAT   | 432  |
| • | Qy . | 772   | GGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGG   | 831  |
|   | Db   | 433   | GGAGGCTGGTCCAGCTGGGCAGAGTGGTCACCCTGTTCCAATCGCTGTGGCCGAGGCTGG   | 492  |
|   | Qy   | 832   | CAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAG   | 891  |
|   | Db   | 493   | CAGAAGCGTACTCGGACCTGCACCAATCCAGCCCCACTCAATGGAGGCGCCTTCTGTGAG   | 552  |
|   | Qy   | 892   | GGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGC   | 951  |
|   | Db   | 553   | GGACAGGCCTTCCAGAAGACAGCTTGCACCACCGTGTGCCCAGTGGATGGA  | 612  |
|   | Qу   | 952   | CCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGC   | 1011 |
|   | Db   | 613   | GAGTGGAGCAAGTGGTCTGCCTGCAGCACAGAGTGTGCGCACTGGCGCAGCCGCGAGTGC   | 672  |
|   | Qy 1 | 1012  | TCTGACCCAGCACCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGC  | 1071 |
|   | Db   | 673   | ATGGCACCGCCACCCCAGAACGGAGGCCGTGACTGCAGCGGGACGCTACTTGACTCCAAG   | 732  |
|   | Qy 1 | 1072  | AACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGAC   | 1119 |
|   | Db   | 733   | AACTGCACTGATGGGCTGTGCGTGCTGAATCAGAGAACTCTAAACGACCCTAAAAGCCAC   | 792  |
|   | Ov : | 1120  | GTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGC  | 1158 |

|   | Db       | 793  |   | 852  |
|---|----------|------|---|------|
|   | Qy<br>Db |      | CTGGTCCTGCTGCTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTG               |      |
|   | Qу       |      | GACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATC           |      |
|   | Db .     |      |   |      |
|   | Qy       | 1273 | AAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGC              | 1326 |
|   | Db       | 973  |   | 1032 |
|   | QУ       | 1327 | ACCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGG                     | 1373 |
|   | Db       | 1033 | GCCAGTGCTGCATCTACCGCGGGCCTGTGTATGCCCTGCAGGACTCCGCCGACAAGATC         | 1092 |
|   | Qу       | 1374 |   | 1401 |
|   | Db       | 1093 | CCCATGACTAATTCGCCCCTGCTGGATCCCCTGCCCAGCCTCAAGATCAAGGTCTATAAC        | 1152 |
|   | Qy       | 1402 | GGGCACCTGCTCAGCCCCCTGGGTGGCGGC                                      | 1431 |
|   | Db       | 1153 | ${\tt TCCAGCACCATCGGTTCTGGGTCTGGCCTGGTGATGGAGCCGACCTGCTGGGTGTCCTC}$ | 1212 |
|   | Qу       |      | CGCCACACACTGCACCACCTCTCAGGCCGAGGAGTTCGTCTCCCGCCTC                   |      |
|   | Db       |      | CCGCCGGGCACGTACCCAGGCGATTTCTCCCGGGACACCCATTTCCTGCACCTGCGCAGT        |      |
|   | Qy<br>Db |      | TCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT              |      |
|   | Qy       |      | GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC        |      |
|   | Db       |      |   |      |
|   | Qу       | 1606 | ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG        | 1665 |
|   | Db       | 1393 |   | 1452 |
|   | Qy       | 1666 | CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC        | 1725 |
| • | Db       | 1453 | GCCGAAAGCACCCTCCCACTTTCAGAAGGTTCCCAGACAGTATTGAGCCCCTCGGTGACC        | 1512 |
|   | Qy       | 1726 | TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG         | 1785 |
|   | Db       | 1513 | TGTGGGCCCACAGGCCTACTCCTGTGCCGCCCTGTCGTCCTCACCGTGCCCCACTGTGCT        | 1572 |
|   | Qy       | 1786 | GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG        | 1845 |
|   | Db       | 1573 | GAAGTCATCGCTGGAGACTGGATCTTTCAGCTCAAGACCCAGGCCCATCAGGGCCACTGG        | 1632 |
|   | Qу       | 1846 | GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG        |      |

| Db   | 1633 | GAGGAGGTGGTGACCTTGGATGAGGAGACCCTCAACACCCTGCTACTGCCAGCTGGAG    | 1692 |
|------|------|---|------|
| Qy   | 1906 | GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC  | 1965 |
| Db   | 1693 | GCTAAGTCCTGCCACATCCTGGTGGACCAGCTGGGTACCTACGTATTCATGGGCGAGTCC  | 1752 |
| Qу   | 1966 | CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC   | 2025 |
| Db   | 1753 | TACTCTCGCTCTGCAGTCAAGCGGCTCCAGCTGGCCATCTTCGCCCCAGCCCTCTGCACC  | 1812 |
| Qy   | 2026 | TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG  | 2085 |
| Db   | 1813 | TCCCTGGAGTATAGCCTCAGGGTCTACTGTCTGGAGGACACCCTGTAGCACTGAAGGAG   | 1872 |
| Qy   | 2086 | GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC | 2145 |
| Db   | 1873 | GTCCTGGAGCTGGAGAGCCTCTGGGTGGCTACTTGGTGGAGGAGCCCAAGCCTTTGCTC   | 1932 |
| Qy   | 2146 | TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG  | 2205 |
| Db   | 1933 | TTTAAGGACAGTTACCACAACCTACGCCTCTCCCTCCATGACATCCCCCATGCCCACTGG  | 1992 |
| Qу   | 2206 | AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG  | 2265 |
| Db   | 1993 | AGGAGCAAACTACTGGCCAAGTACCAGGAGATTCCCTTCTACCACGTCTGGAATGGCAGC  | 2052 |
| Qу   | 2266 | CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG  | 2325 |
| Db   | 2053 | CAGAGAGCCCTGCACTGCACTTTCACCCTGGAGAGGCATAGCCTGGCCTCCACGGAGTTC  | 2112 |
| Qу   | 2326 | GCCTGCAAGCTGTGGCTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC  | 2385 |
| Dp . |      | ACCTGTAAGGTCTGCGTGCGGCAGGTCGAAGGGGGAAGGCCAGATTTTCCAGCTGCACACA |      |
| QУ   |      | AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGG     |      |
| Db   | 2173 | ACGTTGGCCGAGACGCCTGCTGGCTCCCTGGATGCTCTCTGCTCTGCCCCGGGCAATGCC  | 2232 |
| Qy   |      | GTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATA  | 1    |
| Db   |      | ATCACCACCCAGCTGGGACCCTATGCCTTCAAGATACCCCTGTCCATCCGCCAAAAGATC  |      |
| Qу   | •    | ATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAA  |      |
| Db . |      | TGCAGCAGCCTGGACGCCCCAACTCCCGGGGCAACGACTGGAGGCTGTTGGCGCAGAAG   |      |
| Qу   |      | CTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATC   |      |
| Db   |      | CTGTCCATGGACCGGTACCTAAACTACTTCGCCACCAAAGCTAGTCCCACAGGTGTCATC  |      |
| QУ   |      | CTCAACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC              |      |
| Db   |      | TTAGACCTCTGGGAAGCTCGGCAACAGGATGACGGGGACCTCAACAGCCTGGCCAGTGCC  |      |
| QУ   |      | GTGGCTGGACTGGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA       |      |
| Db   | 2473 | TTGGAGGAGATGGGCAAGAGTGAGATGCTGGTAGCCATGGCCACAGATGGCGATTGCTGA  | 2532 |

```
BI758231
LOCUS
           BI758231
                                   1034 bp
                                             mRNA
                                                     linear
                                                              EST 25-SEP-2001
DEFINITION
           603029876F1 NIH MGC 114 Homo sapiens cDNA clone IMAGE:5200171 5',
           mRNA sequence.
ACCESSION
           BI758231
           BI758231.1 GI:15749809
VERSION
KEYWORDS
           EST.
SOURCE
           Homo sapiens (human)
 ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
           1 (bases 1 to 1034)
REFERENCE
 AUTHORS
           NIH-MGC http://mgc.nci.nih.gov/.
           National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE
 JOURNAL
           Unpublished (1999)
           Contact: Robert Strausberg, Ph.D.
COMMENT
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM11501 row: g column: 20
           High quality sequence stop: 793.
FEATURES
                    Location/Qualifiers
                    1. .1034
    source
                    /organism="Homo sapiens"
                    /mol type="mRNA"
                    /db xref="taxon:9606"
                    /clone="IMAGE:5200171"
                    /lab_host="DH10B"
                    /clone lib="NIH MGC 114"
                    /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
                    Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
                    male brains, age range 23-27 yo. Library is oligo-dT
                    primed and directionally cloned (EcoRV site is destroyed
                    upon cloning). Average insert size 1.5 kb, insert size
                    range 1-3 kb. Library is normalized and enriched for
                    full-length clones and was constructed by C. Gruber
                    (Invitrogen). Research Genetics tracking code 019. Note:
                    this is a NIH MGC Library."
ORIGIN
                         27.9%; Score 768.6; DB 4;
                                                     Length 1034;
  Query Match
                         91.2%; Pred. No. 1.6e-151;
 Best Local Similarity
                                0; Mismatches
                                                74;
                                                     Indels
                                                              15; Gaps
                                                                           9;
  Matches 918; Conservative
         150 GCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACAT 209
Qy
             1 GCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACAT 60
Db
         210 CGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTT 269
Qу
```

RESULT 9

| Db   | 61   | CGTCAAGAACAAGCCAGTGCTGCTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTT  | 120  |
|------|------|--|------|
| Qу   | 270  | CAAGTGCAACGGGGAGTGGGTGCGCCAGGTGACCACGTGATCGAGCGCACACAGACGG   | 329  |
| Db   | 121  | CAAGTGCAACGGGGAGTGGGTGCCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGG  | 180  |
| Qу   | 330  | GAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAA | 389  |
| Db   | 181  | GAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAA | 240  |
| Qу   | 390  | GGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCAC | 449  |
| Db   | 241  | GGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCAC | 300  |
| Qy . | 450  | CAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCC | 509  |
| Db   | 301  | CAAGAGTCAGAAGGCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCC | 360  |
| Qу   | 510  | GCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGG | 569  |
| Db   | 361  | GCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCCCTGCCGTCCACCGGAGGG    | 420  |
| Qу   | 570  | CATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGA | 629  |
| Db   | 421  | CATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGA | 480  |
| Qу   | 630  | CCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGA | 689  |
| Db   | 481  | CCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGA | 540  |
| Qу   | 690  | CACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGC | 749  |
| Db   | 541  | CACGGACAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGC | 600  |
| Qу   | 750  | TGCTGTCATCGTCAACGGTGGGTGGTCGA-CGTGGACCGAGTGGTCCGTCTGCA       | 808  |
| Db   | 601  | TGCTGTCATCGTCAACGGTGGGTGGTCGACCGTGGACCGAGTGGTCCGTCTGC-       | 659  |
| Qу   | 809  | GCGCCAGCTGTGGGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTC   | 868  |
| Db   | 660  | GCGCCAGCTGTGGGCGGCTGGCAGAAACCGGAGCCGGAGCTGCACAACCCGGTGCCTC   | 719  |
| Qу   | 869  | TCAACGGGGGCGCTTTCTGTGA-GGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTG | 927  |
| Db   | 720  | TCAACGGGGGCGCTTTCTGTGAGGGGGCAGAATGTCCAGAAAGCAGCTGCGCCACCT    | 776  |
| Qу   | 928  | TGCCCAGTAGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACT   | 985  |
| Db   | 777  | GTGCCAGTGGACGCAGCTGTAGCACGTGGAGCCAGTGGTCGGCCTGTTGGGCTT       | 836  |
| Qу   | 986  | GCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAG      | 1043 |
| Db   | 837  | GCACCCACTTGGGCGGAGCCGGAGTGCTCTGAACCCAGCACCCCGGCACGGCAGGGGGAG | 896  |
| Qу   | 1044 | GTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAG    | 1100 |

```
Db
          897 GTGTGCCCAGGTCACTGGACCTGGCACCCGGGA-TGGTCCAGTGAGCTCTGTGT-CCCAC 954
         1101 TGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCG 1147
Ov
               955 GGGTTCTGGCCCTGAGGACTTGGCCTCCTATGTGGGCCTCATCCCGG 1001
Db
RESULT 10
BC033727
LOCUS
                                     1532 bp . mRNA
                                                        linear
                                                                 HTC 25-MAR-2004
            Homo sapiens unc-5 homolog A (C. elegans), mRNA (cDNA clone
DEFINITION
            IMAGE: 5166762), containing frame-shift errors.
ACCESSION
VERSION
            BC033727.1 GI:21707230
KEYWORDS
            HTC.
SOURCE
            Homo sapiens (human)
  ORGANISM
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
               (bases 1 to 1532)
  AUTHORS
            Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
            Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
            Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
            Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
            Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
            Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
            Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
            Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
            Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
            McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
            Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
            Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
            Fahey, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S.,
            Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
            Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
            Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
            Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
            Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
  TITLE
            Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
            Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  JOURNAL
            12477932
   PUBMED
               (bases 1 to 1532)
REFERENCE
  AUTHORS
            Strausberg, R.
  TITLE
            Direct Submission
  JOURNAL
            Submitted (02-JUL-2002) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            NIH-MGC Project URL: http://mgc.nci.nih.gov
  REMARK
COMMENT
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
```

Sequencing Center (NISC),

```
Web site: http://www.nisc.nih.gov/
          Contact: nisc mgc@nhgri.nih.gov
          Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
          Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
          Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
          Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
          Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
          McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
          Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
          Young, A., Zhang, L.-H. and Green, E.D.
          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
          Series: IRAK Plate: 68 Row: i Column: 2
          This clone was selected for full length sequencing because it
          passed the following selection criteria: GenomeScan gene prediction
          This clone has the following problem: frame shifted.
FEATURES
                  Location/Qualifiers
                  1. .1532
    source
                  /organism="Homo sapiens"
                  /mol type="mRNA"
                  /db xref="taxon:9606"
                  /clone="IMAGE:5166762"
                  /tissue type="Brain, adult medulla"
                  /clone lib="NIH MGC 119"
                  /lab host="DH10B"
                  /note="Vector: pCMV-SPORT6"
ORIGIN
 Query Match
                       26.8%;
                              Score 736.6; DB 3;
                                                Length 1532;
 Best Local Similarity
                       84.2%;
                              Pred. No. 9.7e-145;
 Matches 917; Conservative
                             0; Mismatches
                                             4;
                                                Indels
                                                       168;
                                                             Gaps
                                                                    1;
        337 GGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTC 396
Qy
            Db
        612 GGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTC 671
        397 GGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGT 456
Qy
            672 GGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGT 731
Db
         457 CAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCC 516
Qy
                                    11111111
        732 CAGAAGGCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCC 791
Db
Qу
         517 AAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCT 576
            792 AAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCT 851
Db
        577 CCAGCCGAGGTGGACTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAAT 636
Qу
            Db
         852 CCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAAT 911
Qy
        637 GTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCC 696
            Db
         912 GTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCC 971
```

Gaithersburg, Maryland;

| Qу | 697  | AACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTC         | 756  |
|----|------|--|------|
| Db | 972  | AACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTC         | 1031 |
| Qу |      | ATCGTCTACGTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGC             |      |
| Db | 1032 | ATCGTCTACGTG   | 1043 |
| Qу | 817  | TGTGGGCGCGGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGG          | 876  |
| Db | 1044 |  | 1043 |
| Qу | 877  | ${\tt GGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTA}$ | 936  |
| Db | 1044 |  | 1043 |
| Qу | 937  | GACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGG         | 996  |
| Db | 1044 | GACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGG         | 1103 |
| Qу | 997  |  | 1056 |
| Db | 1104 | CGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACT         | 1163 |
| Qу | 1057 | GACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAG         | 1116 |
| Db | 1164 | GACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAG         | 1223 |
| Qу | 1117 | GACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTT            | 1176 |
| Db | 1224 |  | 1283 |
| Qy | 1177 | GTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCG         | 1236 |
| Db | 1284 | GTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCG         | 1343 |
| Qy | 1237 | TCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCC         | 1296 |
| Db | 1344 | TCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCC         | 1403 |
| ДĀ | 1297 | CATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCAGGGCAGTCTC          | 1356 |
| Db | 1404 | CATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCAGGGCAGTCTC          | 1463 |
| Qy | 1357 | TGTCCCCGGCAGGATGGGCCCCAAGCTCCAGCTCACCAATGGGCACCTGCTCAGC              | 1416 |
| Db | 1464 | TGTCCCCGGCAGGATGGGCCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGC        | 1523 |
| Qу | 1417 | CCCCTGGGT 1425   |      |
| Db | 1524 |  |      |

RESULT 11 AI951556

```
wv36f04.x1 NCI CGAP Ov18 Homo sapiens cDNA clone IMAGE:2531647 3'
DEFINITION
           similar to TR:008721 008721 TRANSMEMBRANE RECEPTOR UNC5H1. ;, mRNA
           sequence.
ACCESSION
           AI951556
           AI951556.1 GI:5743866
VERSION
KEYWORDS
           EST.
SOURCE
           Homo sapiens (human)
 ORGANISM
           Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              (bases 1 to 788)
REFERENCE
           NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 AUTHORS
 TITLE
           National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
           Tumor Gene Index
  JOURNAL
           Unpublished (1997)
           Contact: Robert Strausberg, Ph.D.
COMMENT
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
           R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento
           Soares, Ph.D. cDNA Library Arrayed by: Christa Prange, The
           I.M.A.G.E. Consortium DNA Sequencing by: Washington University
           Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           www-bio.llnl.gov/bbrp/image/image.html
           Insert Length: 1125
                               Std Error: 0.00
           Seq primer: -40UP from Gibco
           High quality sequence stop: 446.
FEATURES
                    Location/Qualifiers
                    1. .788
   source
                    /organism="Homo sapiens"
                    /mol type="mRNA"
                    /db xref="taxon:9606"
                    /clone="IMAGE: 2531647"
                    /tissue type="fibrotheoma"
                    /lab host="DH10B (phage-resistant)"
                    /clone lib="NCI CGAP Ov18"
                    /note="Organ: ovary; Vector: pT7T3D-Pac (Pharmacia) with a
                    modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
                    strand cDNA was primed with a Not I - oligo(dT) primer [5'
                    double-stranded cDNA was ligated to Eco RI adaptors
                    (Pharmacia), digested with Not I and cloned into the Not
                    I and Eco RI sites of the modified pT7T3 vector. Library
                    went through one round of normalization, and was
                    constructed by Bento Soares and M. Fatima Bonaldo.
ORIGIN
                        26.7%;
                                Score 735.2; DB 1;
                                                    Length 788;
  Query Match
                        96.5%; Pred. No. 1.8e-144;
  Best Local Similarity
 Matches 749; Conservative
                               0; Mismatches
                                               27;
                                                    Indels
                                                              0;
                                                                  Gaps
        1469 CCGAGGAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCA 1528
Qy
             1 CCGAGGAGTTCGTCTCCGCCTCTCCGCCCAGAACTACTTCCGCTCCCTGCCCCGAGGCA 60
Db
        1529 CCAGCAACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATA 1588
```

Qy

| Db | 61   |   | 120  |
|----|------|---|------|
| Qy | 1589 | CAGGTATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCT   | 1648 |
| Db | 121  | CAGGAATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCT   | 180  |
| Qу | 1649 | ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTG             | 1708 |
| Db | 181  | ACCTCACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTG             | 240  |
| Qу | 1709 | TGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTCCTCACCCGGCCAGTCATCCTGG     | 1768 |
| Db | 241  | TGAGTCCCATCGTTAGCTGTGGACCCCCTGGCGTTCTGCTCACCCGGCCAGTCATCCTGG  | 300  |
| Qу | 1769 | CTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGT  | 1828 |
| Db | 301  | CTATGGACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGT  | 360  |
| Qу | 1829 | CGTGCGAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCT  | 1888 |
| Db | 361  |   | 420  |
| Qу | 1889 | ACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTG  | 1948 |
| Db | 421  | ACTACTGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGAGCCGCTATG  | 480  |
| Qу | 1949 | CCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTG  | 2008 |
| Db | 481  | CCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTG  | 540  |
| Qу | 2009 | CGCCGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC           | 2068 |
| Db | 541  | CGCCGGTGGCCTGCACCTCCCTCGAGTACAACATACTGGTCTACTGCCTGC           | 600  |
| Qу | 2069 | ACGATGCACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGG  | 2128 |
| Db | 601  | ACGATGCACTCAACGTAGTGGTGCAGCTGGAGAAGCAGCTGCAGGGACAGCTGATCCAGG  | 660  |
| Qу | 2129 | AGCCACGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC        | 2188 |
| Db | 661  | AGCCACTGGTACTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATNCATCCACGATG  | 720  |
| Qу | 2189 | TGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTT 2244 | 1    |
| Db | 721  | TGCCCAGCTNCCNTGTGGAGAGTAAGCTTCTTGTCAGCTACCCAGAGATCCNCTAT 776  |      |

# RESULT 12 BX348193/c

LOCUS BX348193 796 bp mRNA linear EST 08-APR-2004 DEFINITION BX348193 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens

cDNA clone CS0DB008YE02 5-PRIME, mRNA sequence.

ACCESSION BX348193

VERSION BX348193.2 GI:46286231

KEYWORDS EST.

SOURCE Homo sapiens (human)

```
ORGANISM
          Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
             (bases 1 to 796)
REFERENCE
          Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 AUTHORS
          Full-length cDNA libraries and normalization
 TITLE
          Unpublished (2001)
 JOURNAL
          On May 5, 2003 this sequence version replaced gi:30367258.
COMMENT
          Contact: Genoscope
          Genoscope - Centre National de Sequencage
          2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
          Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
          1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
          end enriched, double-strand cDNA was digested with Not I and cloned
          into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
          was normalized. Library was constructed by Life Technologies, a
          division of Invitrogen. This sequence belongs to sequence cluster
          3239.r
          For more information about this cluster, see
          http://www.genoscope.cns.fr/cdna?s=CSOBAF004ZD01 AF00293 1&c=3239.r
FEATURES
                  Location/Qualifiers
                  1. .796
    source
                  /organism="Homo sapiens"
                  /mol type="mRNA"
                  /db xref="taxon:9606"
                  /clone="CS0DB008YE02"
                  /tissue type="NEUROBLASTOMA COT 10-NORMALIZED"
                  /clone lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"
                  /note="1st strand cDNA was primed with a NotI-oligo(dT)
                  primer. Five prime end enriched, double-strand cDNA was
                  digested with Not I and cloned into the Not I and EcoR V
                  sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
                              Score 721.6; DB 5;
                                                 Length 796;
                       26.2%;
 Best Local Similarity
                       96.4%; Pred. No. 1.3e-141;
 Matches 758; Conservative
                             0; Mismatches
                                            26;
                                                 Indels
                                                          2;
                                                             Gaps
                                                                     2;
        1679 GGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTG 1738
Qy
            795 GTTGCCCCAAACCGGGCTGTCAGACCCTGTTGAGTCCCATGGTTAGCTGTGAA-CCCCTG 737
Db
        1739 GCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTG 1798
Qу
            736 GCGTCCTGCTCACCGGCCCAGTCATCCTGGGTATGGACCACTGTGGGGAGCCCAGCCCTG 677
Db
        1799 ACAGCTGGAGCCTGCGCCTC-AAAAAGCAGTCGTGCGAGGCAGCTGGGAGGATGTGCTG 1857
Qу
                    676 ACAGCTGAGGCCTGCGCCTCAAAAAAACAGTCGTGCGAGGACAGCTGGGAGTATGTCCTG 617
Db
        1858 CACCTGGGCGAGGGGCGCCCTCCCACCTCTACTGCCAGCTGGAGGCCAGTGCCTGC 1917
Qу
            616 CACCTGGGCGAGNAGGCGCCCTCCCACCTCTACTGCCAGCTGGAGGCCAGTGCCTGC 557
Db
        1918 TACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCT 1977
Qу
```

| Db · 5               | 66 TACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCT 497              |
|----------------------|--|
| Qy 19                | 78 GCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC                        |
| Db 4                 | 96 GCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC                        |
| Qy 20                | 38 AACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTG 2097             |
| Db 4                 | 36 AACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTG 377              |
| Qy 20                | 98 GAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGT 2157             |
| Db 3                 | 76 GAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGT 317              |
| Qy 21                | 58 TACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTC 2217             |
| Db 3                 | 16 TACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTC 257              |
| Qy 22                | 18 CTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTG 2277             |
| Db 2                 | 66 CTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTG 197              |
| Qy 22                | 78 CACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTG 2337             |
| Db 1                 | 96 CACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTG 137              |
| Qy 23                | 38 TGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAG 2397             |
| Db . 1               | 36 TGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACATCACCAAG 77                     |
| Qy 23                | 98 GACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTG 2457             |
| Db                   | 76 GACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGNCCCTGGT 17               |
| Qy 24                | 58 GGCCCC 2463   |
| Dp.                  | 16 GGGCCC 11   |
| RESULT 13            |  |
| BI818609<br>LOCUS    | BI818609 818 bp mRNA linear EST 04-OCT-2001                                      |
| DEFINITION           | 603033362F1 NIH MGC 115 Homo sapiens cDNA clone IMAGE:5174559 5', mRNA sequence. |
| ACCESSION            | BI818609   |
| VERSION<br>KEYWORDS  | BI818609.1 GI:15929902<br>EST.   |
| SOURCE               | Homo sapiens (human)   |
| ORGANISM             | Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;   |
|                      | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.                       |
| REFERENCE<br>AUTHORS | 1 (bases 1 to 818) NIH-MGC http://mgc.nci.nih.gov/.                              |
| TITLE                | National Institutes of Health, Mammalian Gene Collection (MGC)                   |
| JOURNAL<br>COMMENT   | Unpublished (1999) Contact: Robert Strausberg, Ph.D.                             |
|                      | Email: cgapbs-r@mail.nih.gov   |

```
Tissue Procurement: Life Technologies, Inc.
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM11434 row: 1 column: 16
          High quality sequence stop: 744.
                  Location/Qualifiers
FEATURES
                  1. .818
    source
                  /organism="Homo sapiens"
                  /mol type="mRNA"
                  /db xref="taxon:9606"
                  /clone="IMAGE:5174559"
                  /lab host="DH10B"
                  /clone lib="NIH MGC 115"
                  /note="Organ: pooled brain, lung, testis; Vector:
                  pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
                  source anonymous pool of 6 male brains, age range 23-27; 1
                  male lung, age 27; and 1 male testis, age 69. Library is
                  oligo-dT primed and directionally cloned (EcoRV site is
                  destroyed upon cloning). Average insert size 1.8 kb,
                  insert size range 1-3 kb. Library is normalized and
                  enriched for full-length clones and was constructed by C.
                  Gruber (Invitrogen). Research Genetics tracking code
                  021. Note: this is a NIH MGC Library."
ORIGIN
 Query Match
                      26.0%;
                             Score 716.4; DB 4;
                                               Length 818;
                      97.7%;
                             Pred. No. 1.7e-140;
 Best Local Similarity
                            0; Mismatches
 Matches 748; Conservative
                                           16;
                                               Indels
                                                        2;
                                                                   2;
         43 GCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGG 102
Qy
            1 GCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGG 60
Db
        103 CTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAAC 162
Qy
            61 CTCCGCGGCTCGGCTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAAC 120
Db
        163 CCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAG 222
Qy
            121 CCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAG 180
Db
        223 CCAGTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGG 282
Qy
            181 CCAGTGCTGCTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGG 240
Db
        283 GAGTGGGTGCCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTG 342
Qу
            241 GAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCACACAGACGGGAGCAGTGGGCTG 300
Db
```

343 CCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTG 402

301 CCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTG 360

Qу

Db

| Db Qy Db Qy Db Qy Db   |
|--|
| Db<br>Qy<br>Db   |
| Qy<br>Db<br>Qy   |
| Db<br>Qy   |
| Qy   |
| <u>-</u>   |
| Db   |
|  |
| Qу   |
| Db   |
| Qу   |
| Db   |
| Qy   |
| Db   |
| RESULT 14 BX364574 LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM REFERENCE AUTHORS |

```
http://www.genoscope.cns.fr/cdna?s=CS1AC006ZF06QP1&c=3529.f.
               Location/Qualifiers.
FEATURES
               1. .853
   source
               /organism="Homo sapiens"
               /mol type="mRNA"
               /db xref="taxon:9606"
               /clone="CS0DC023YB10"
               /tissue type="NEUROBLASTOMA COT 25-NORMALIZED"
               /clone lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
               /note="1st strand cDNA was primed with a NotI-oligo(dT)
               primer. Five prime end enriched, double-strand cDNA was
               digested with Not I and cloned into the Not I and EcoR V
               sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
 Query Match
                   25.1%;
                         Score 692; DB 5; Length 853;
 Best Local Similarity
                   99.0%;
                        Pred. No. 2.4e-135;
Matches 703; Conservative
                        3; Mismatches
                                      3;
                                         Indels
                                                 1;
                                                          1;
                                                    Gaps
         1 CCGCGGGGCCCGCCCGCCCGCCCGCCCGCCGCGCCATGGCCGTCCGGCCC 60
Qу
          Db
        61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120
Qу
          205 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 264
Db
       121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180
Qу
          265 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCTGGTGCYAACCCGGACCTGCTTCCCCAC 324
Db
       181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240
Qy
          325 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 384
Db
       241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300
Qу
          385 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 444
Db
       301 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC 360
Qу
          445 GACCACGTGATCGAGCGCA-SACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC 503
Db
       361 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 420
Qу
          504 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 563
Db
       421 TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 480
Qу
          564 TGCGTGGCATGGAGCTCCTCGGGCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 623
Db
       481 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 540
Qу
            624 TATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 683
Db
       541 ATCGTGCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 600
Qy
```

```
Db
         684 ATCGTGCTGCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 743
         601 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 660
Qy
             744 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 803
Db
         661 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGT 710
Qу
             Db
         804 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGT 853
RESULT 15
AY411748
                                                             GSS 12-DEC-2003
LOCUS
           AY411748
                                  2532 bp
                                            DNA
                                                    linear
           Pan troglodytes HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION
           genomic survey sequence.
ACCESSION
           AY411748
           AY411748.1 GI:39767716
VERSION
KEYWORDS
           GSS.
SOURCE
           Pan troglodytes (chimpanzee)
           Pan troglodytes
  ORGANISM
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
             (bases 1 to 2532)
 AUTHORS
           Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
           Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
           Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
           Adams, M.D. and Cargill, M.
  TITLE
           Inferring nonneutral evolution from human-chimp-mouse orthologous
           gene trios
           Science 302 (5652), 1960-1963 (2003)
  JOURNAL
  PUBMED
           14671302
              (bases 1 to 2532)
REFERENCE
           2
           Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
  AUTHORS
           Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
           Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
           Adams, M.D. and Cargill, M.
           Direct Submission
  TITLE
  JOURNAL
           Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
           Rockville, MD 20850, USA
           This sequence was made by sequencing genomic exons and ordering
COMMENT
           them based on alignment.
FEATURES
                    Location/Oualifiers
                    1. .2532
    source
                    /organism="Pan troglodytes"
                    /mol type="genomic DNA"
                    /db xref="taxon:9598"
                    <1. .>2532
    gene
                    /locus tag="HCM4327"
ORIGIN
                                Score 678.2; DB 9;
 Query Match
                        24.6%;
                                                    Length 2532;
  Best Local Similarity
                        54.4%;
                                Pred. No. 2.2e-132;
                               0; Mismatches 1021;
 Matches 1370; Conservative
                                                    Indels 129; Gaps
                                                                         9;
Qу
         352 GAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATAC 411
```

| Db   | 13   | GAGGTGCAGATCGAGGTGTCGCGGCAGCAGGTGGAGGAGCTCTTTGGGCTGGAGGATTAC | 72   |
|------|------|--|------|
| Qy   | 412  | TGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATC | 471  |
| Db   | 73   | TGGTGCCAGTGCGTGGCCTGGAGCTCTGCGGGCACCACCAAGAGTCGCCGAGCCTACGTC | 132  |
| Qу   |      | CGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTG | 531  |
| Db   |      | CGCATCGCCTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN                     | 192  |
| Qу   | 532  | GAGCAGGGCATCGTGCCGTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAG    | 591  |
| Db   | 193  | NUNNNUNNNUNNNNNNNNNNNNNNNNNNNNNNNNNNNNN                      | 252  |
| Qу   | 592  | TGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGG | 651  |
| Db   | 253  | TGGCTCAAGAATGAGGATGTCATCGACCCCACCCAGGACACCAACTTCCTGCTCACCATC | 312  |
| Qу   | 652  | GAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTG | 711  |
| Db . | 313  | GACCACAACCTCATCCGCCAGGCCCGCCTGTCGGACACTGCCAACTATACCTGCGTG    | 372  |
| Qу   | 712  | GCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAAC | 771  |
| Db   | 373  | GCCAAGAACATCGTGGCCAAACGCCGGAGCACCACTGCCACCGTCATCGTCTACGTGAAT | 432  |
| Qу   | 772  | GGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGG | 831  |
| Db   | 433  | GGCGGCTGGTCCAGCTGGGCAGAGTGGTCGCCCTGCTCCAACCGCTGTGGNNGANNCTGG | 492  |
| Qу   | 832  | CAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAG | 891  |
| Db   | 493  | CAGAAGCGCACCCGGACCTGCACCAACCCCGCCCACTCAACGGAGGGNNNNTCTGCGAG  | 552  |
| Qy   | 892  | GGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGC | 951  |
| Db   | 553  | GGCCANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN                      | 612  |
| Qy   | 952  | CCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGC | 1011 |
| Db   | 613  | GAGTGGAGCAAGTGGTCAGCCTGCAGCACTGAGTGTGCCCACTGGCGTAGCCGCGAGTGC | 672  |
| Qy   | 1012 | TCTGACCCAGCACCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGC  | 1071 |
| Db   | 673  | ATGGCGCCCCAÇCCCAGAACGGAGGCCGTGACTGCAGCGGGACGCTGCTCGACTCTAAG  | 732  |
| Qy   | 1072 | AACTGTACCAGTGACCTCTGTGTACACAGTGCTTCT                         | 1107 |
| Db   | 733  | AACTGCACAGATGGGCTGTGCATGCAAAATAAGAAAACTCTAAGCGACCCCAACAGCCAC | 792  |
| Qy   | 1108 | GGCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTC             | 1155 |
| Db   | 793  | CTGCTGGAGGCCTCAGGGGATGCGGCGCTGTATGCGGGGCTCGTGGTGGCCGTCTTCNTG | 852  |
| Qy   | 1156 | TGCCTGGTCCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTG    | 1215 |
| Db   | 853  | GTCNTGGCAATCCTCATGGCGGTGGGGGTGGTGTTACCGCCGCAACTGCCGTGACTTC   | 912  |

|   | QУ   |        | ACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATC     | 1272 |
|---|------|--------|--|------|
|   | Db   | 913 G  | CACAGACATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTTT   | 972  |
|   | Qу   |        | AGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGACCTCAGC        | 1326 |
| • | Db   | •      | AGACGGCAAGGCCCAGCAACCCGCAGCTCCTACACCCCTCTGTGCCTCCTGACCTGACA  | 1032 |
|   | Qy   |        | CCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGG              | 1374 |
|   | Db   |        | CCAGCGCCGGCATCTACCGCGGACCCGTGTATGCCCTGCAGGACTCCACCGACAAAATC  | 1092 |
|   | Qу   | 1375 - | CCCAGCCCCAAGTTCCAGCTCACCA                                    | 1399 |
|   | Db   | 1093 C | CCATGACCAACTCTCCTCTGCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGC  | 1152 |
|   | Qу   | 1400 - | ATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACA                    | 1440 |
|   | Db   | 1153 Т | CCAGCACTACGGGCTCTGGGCCAGGCCTGGCAGATGGGGCTGACCTGCTGGGAGTCTTG  | 1212 |
|   | Qу   |        | TGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG   | 1500 |
|   | Db   | •      | CGCCTGGCACATACCCTAGCGATTTCGCCCGGGACACCCACTTCCTGCACCTGCGCAGC  | 1272 |
|   | Qу   | 1501 A | ACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT                 | 1545 |
|   | Db   | 1273 G | CCAGCCTCGGTTCCCAGCAGCTCTTGGGCCTGCCCCGAGACCCAGNNAGCAGCGTCAGC  | 1332 |
|   | Qy ' |        | GGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC  | 1605 |
|   | Db   |        | GCACCTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN                     | 1392 |
|   | Qy   |        | TCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG  | 1665 |
|   | Db   |        | TGCCCAATGGAGCCATTCCCCAGGGCAAGTTCTACGAGATGTATCTACTCATCAACAAG  | 1452 |
|   | Qy   |        | CGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC  | 1725 |
|   | Db   |        | CAGAAAGTACCCTCCCGCTTTCAGAAGGGACCCAGACAGTATTGAGCCCCTCAGTGACC  | 1512 |
|   | Qy   |        | *GTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG  | 1785 |
|   | Db   |        | 'GTGGACCCACAGGCCTCCTGTGCCACCCCGTCATCCTCACCATGCCCCACTGTGCC    | 1572 |
|   | Qу   |        | AGCCCAGCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG   | 1845 |
|   | Db   | •      | CAAGTCAGTGCCCGTGACTGGATCTTTCAGCTCAAGACCCAGGCCCACCAGGGCCACTGG | 1632 |
|   | Qу   |        | SAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG | 1905 |
| • | Db   |        | BAGGAGGTGGTGACCCTGGATGAGGAGACCCTGAACACCCCTGCTACTGCCAGCTGGAG  | 1692 |
|   | Qy . |        | CCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC  | 1965 |
|   | Db   |        |  | 1752 |

| Qу | 1966 | CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC  | 2025 |
|----|------|--|------|
| Db | 1753 | TATTCCCNNNNNNCAGTCNNNNNGCTCCNGCTGGCCGTCTTNGCCCCTGCCCTCTGCACC | 1812 |
| Qу | 2026 | TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG | 2085 |
| Db | 1813 | TCCCTGGANNNNNGCCTCCGGGTCTACTGCCTGGNNNNNNNNNN                 | 1872 |
| Qу | 2086 | GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC | 2145 |
| Db | 1873 | GTGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGGAGGAGCCGAAACCGCTAATG | 1932 |
| Qу | 2146 | TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG | 2205 |
| Db | 1933 | TTCAAGGACAGTTACCACAACCTGCGCCTCTCCCTCCATGACCTCCCCCATGCCCATTGG | 1992 |
| Qу | 2206 | AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG | 2265 |
| Db | 1993 | AGGAGCAAGCTGCTGGCCAAATACCAGGAGATCCCCTTCTGTCACATTTGGAGTGGCAGC | 2052 |
| Qу | 2266 | CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG | 2325 |
| Db | 2053 | CAGAAGGCCCTCCACTGCACTTTCACCCTGGAGAGGCACAGCTTGGCCTCCACAGAGCTC | 2112 |
| Qу | 2326 | GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC | 2385 |
| Db | 2113 | ACCTGCAAGATCTGCGTGCGGCAAGTGGAAGGGGAGGGCCAGATATTCCAGCTGCATACC | 2172 |
| Qy | 2386 | AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGG    | 2442 |
| Db | 2173 | ACTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCCTGGCAGCACT | 2232 |
| Qy | 2443 | GTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATA | 2502 |
| Db | 2233 | GTCACCACCCAGCTGGGACCTTATGCCTTCAAGATCCCACTGTCCATCCGCCAGAAGATA | 2292 |
| Qу | 2503 | ATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAA | 2562 |
| Db | 2293 | TGCAACAGCCTAGATGCCCCCAACTCACGGGGCAATGACTGGCGGATGTTAGCACAGAAG | 2352 |
| Qу | 2563 | CTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATC  | 2622 |
| Db | 2353 | CTCTCTATGGACCGGTACCTGAATTACTTTGCCACCAAAGCGAGCCCCACGGGTGTGATC | 2412 |
| QУ | 2623 | CTCAACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC             | 2682 |
| Db | 2413 | CTGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGGACCTCAACAGCCTGGCGAGTGCC | 2472 |
| Qy | 2683 | GTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA | 2742 |
| Db | 2473 | TTGGAGGAGATGGCAAGAGTGAGATGCTGGTGGCCTGTGGCCACCGACGGGGACTGCTGA | 2532 |

Search completed: March 6, 2005, 10:10:18
Job time: 8663.89 secs

### GenCore version 5.1.6 Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2005, 15:11:26; Search time 11689.8 Seconds

(without alignments)

11407.261 Million cell updates/sec

Title: US-10-624-932-1

Perfect score: 2752

Sequence: 1 ccgcqqqqccccqcqcccqq.....tgaqtqctqaqgccgqccag 2752

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:\*

1: gb ba:\*

2: gb\_htg:\*

3: gb\_in:\*

4: qb om:\*

5: gb\_ov:\*

6: gb pat:\*

7: gb ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb sy:\*

13: gb\_un:\*

14: gb\_vi:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result<br>No. | Score  | %<br>Query<br>Match | Length | DB | ID       | Description       |
|---------------|--------|---------------------|--------|----|----------|-------------------|
| 1             | 2752   | 100.0               | 2752   | 6  | AX449572 | AX449572 Sequence |
| 2             | 2687.4 | 97.7                | 2697   | 6  | AX451652 | AX451652 Sequence |
| 3             | 2676.4 | 97.3                | 2881   | 6  | AX527916 | AX527916 Sequence |

|    |        |      |      | _   |           |                    |
|----|--------|------|------|-----|-----------|--------------------|
| 4  | 2435.4 | 88.5 | 2784 | 6   | CQ730306  | CQ730306 Sequence  |
| 5  | 2356   | 85.6 | 3580 | 6   | AX367094  | AX367094 Sequence  |
| 6  | 2323.2 | 84.4 | 3992 | 10  |           | AJ487852 Mus muscu |
| 7  | 2259   | 82.1 | 3014 | 6   | BD057524  | BD057524 Netrin re |
| 8  | 2252.2 | 81.8 | 2697 | 6   | AX268596  | AX268596 Sequence  |
| 9  | 2252.2 | 81.8 | 2697 | 10  | RNU87305  | U87305 Rattus norv |
| 10 | 1983.6 | 72.1 | 3844 | 10  | BC058084  | BC058084 Mus muscu |
| 11 | 1635.4 | 59.4 | 2688 | 9   | BC009333  | BC009333 Homo sapi |
| 12 | 1562.4 | 56.8 | 1787 | 6   | BD057525  | BD057525 Netrin re |
| 13 | 1304.2 | 47.4 | 9700 | 6   | AX054976  | AX054976 Sequence  |
| 14 | 992    | 36.0 | 9299 | 10  | MMU72634  | U72634 Mus musculu |
| 15 | 991    | 36.0 | 2962 | 5   | AY187310  | AY187310 Gallus ga |
| 16 | 986    | 35.8 | 9328 | 10  | AB118026  | AB118026 Rattus no |
| 17 | 970    | 35.2 | 3646 | 6   | CQ881060  | CQ881060 Sequence  |
| 18 | 970    | 35.2 | 3646 | 9   | AF055634  | AF055634 Homo sapi |
| 19 | 943.8  | 34.3 | 3770 | 9   | AY126437  | AY126437 Homo sapi |
| 20 | 936.2  | 34.0 | 2860 | 6   | AX686445  | AX686445 Sequence  |
| 21 | 936.2  | 34.0 | 2860 | 6   | AX686447  | AX686447 Sequence  |
| 22 | 913.8  | 33.2 | 3672 | 10  | BC048162  | BC048162 Mus muscu |
| 23 | 913.8  | 33.2 | 3672 | 10  | BC057560  | BC057560 Mus muscu |
| 24 | 913.6  | 33.2 | 2895 | 6   | AX512281  | AX512281 Sequence  |
| 25 | 905.6  | 32.9 | 3933 | 6   | CQ850929  | CQ850929 Sequence  |
| 26 | 905.6  | 32.9 | 3933 | 9   | AK128132  | AK128132 Homo sapi |
| 27 | 905.6  | 32.9 | 3935 | 9   | AB096256  | AB096256 Homo sapi |
| 28 | 904.2  | 32.9 | 2986 | 6   | CQ881052  | CQ881052 Sequence  |
| 29 | 902.4  | 32.8 | 3884 | 6   | AR528525  | AR528525 Sequence  |
| 30 | 902.4  | 32.8 | 3884 | 6   | AX464012  | AX464012 Sequence  |
| 31 | 902.4  | 32.8 | 3884 | 9   | AY358351  | AY358351 Homo sapi |
| 32 | 893.8  | 32.5 | 2995 | 6   | AX497288  | AX497288 Sequence  |
| 33 | 892    | 32.4 | 4294 | 10  | AK122575  | AK122575 Mus muscu |
| 34 | 875.6  | 31.8 | 3788 | 10  | MMU487853 | AJ487853 Mus muscu |
| 35 | 852.4  | 31.0 | 2838 | 10  | RNU87306  | U87306 Rattus norv |
| 36 | 843.6  | 30.7 | 2625 | 6   | CQ721377  | CQ721377 Sequence  |
| 37 | 841.4  | 30.6 | 2831 | 6   | BD057526  | BD057526 Netrin re |
| 38 | 764.8  | 27.8 | 813  | 6   | AX054892  | AX054892 Sequence  |
| 39 | 753.6  | 27.4 | 2230 | 6   | CQ845766  | CQ845766 Sequence  |
| 40 | 753.6  | 27.4 | 2230 | 9   | AK131380  | AK131380 Homo sapi |
| 41 | 751.4  | 27.3 | 4330 | 5   | AY744919  | AY744919 Petromyzo |
| 42 | 717.2  | 26.1 | 2832 | 5   | AY099459  | AY099459 Xenopus 1 |
| 42 | 668.4  | 24.3 | 2612 | 6   | CQ881064  |                    |
| 43 | 623.2  | 24.3 | 2612 | 6   | AX800717  | CQ881064 Sequence  |
| 44 | 623.2  |      | 2868 | -   |           | AX800717 Sequence  |
| 43 | 023.2  | 22.6 | 2000 | 6 . | AX800719  | AX800719 Sequence  |

## ALIGNMENTS

```
RESULT 1
AX449572
LOCUS
                                   2752 bp
           AX449572
                                              DNA
                                                      linear
                                                               PAT 03-JUL-2002
DEFINITION Sequence 1 from Patent W00210216.
ACCESSION
           AX449572
VERSION
           AX449572.1 GI:21698195
KEYWORDS
SOURCE
           Homo sapiens (human)
 ORGANISM Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
REFERENCE
 AUTHORS
        Padigaru, M., Mezes, P., Mishra, V., Burgess, C., Casman, S.,
        Grosse, W.M., Alsobrook, J.P., Lepley, D.M., Gerlach, V.L.,
        Macdougall, J.R. and Smithson, G.
        Proteins and nucleic acids encoding same
 TITLE
 JOURNAL
        Patent: WO 0210216-A 1 07-FEB-2002;
        Curagen Corporation (US)
FEATURES
               Location/Qualifiers
               1. .2752
   source
               /organism="Homo sapiens"
               /mol type="unassigned DNA"
               /db xref="taxon:9606"
ORIGIN
                   100.0%;
                          Score 2752;
                                   DB 6;
 Query Match
                                        Length 2752;
 Best Local Similarity
                   100.0%; Pred. No. 0;
                           Mismatches
 Matches 2752; Conservative
                        0;
                                                         0;
                                         Indels
                                                0;
                                      0;
                                                   Gaps
        Qy
          1 CCGCGGGGCCCGGCCCGGCCCGCCTGCCCGCCGGGCCATGGCCGTCCGGCCC 60
Db
        61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120
Qу
          61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120
Db
       121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180
Qy
          121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180
Db
       181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240
Qу
          181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240
Db
       241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300
Qу
          241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300
Db
       301 GACCACGTGATCGAGCGCAGCAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC 360
Qу
          301 GACCACGTGATCGAGCGCAGCAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC 360
Db
       361 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 420
Qy
          Db
       361 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 420
       421 TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 480
Qу
          421 TGCGTGGCATGGAGCTCCTCGGGCACCACAAGAGTCAGAAGGCCTACATCCGCATAGCC 480
Db
       481 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 540
Qy
          481 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 540
Db
       541 ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 600
Qy
```

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

| Db   | 541  | ATCGTGCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG            | 600  |
|------|------|--|------|
| Qу   | 601  | AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC       | 660  |
| Db   | 601  |  | 660  |
| Qу   | 661  | $\tt CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC$ | 720  |
| Db   | 661  | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC       | 720  |
| Qу   | 721  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG           | 780  |
| Db   | 721  |  | 780  |
| Qy   | 781  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG       | 840  |
| Db   | 781  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGGCGGCTGGCAGAAACGG      | 840  |
| Qy . | 841  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT       | 900  |
| Db   | 841  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT       | 900  |
| Qу   | 901  | GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGGAGC       | 960  |
| Db   | 901  |  | 960  |
| Qу   | 961  | AAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCA       | 1020 |
| Db   | 961  | AAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCA       | 1020 |
| Qу   | 1021 | GCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACC         | 1080 |
| Db   | 1021 |  | 1080 |
| Qу   | 1081 | AGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTC       | 1140 |
| Db   | 1081 | AGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTC       | 1140 |
| Qу   | 1141 | ATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTCATCCTCGTTTATTGCCGG             | 1200 |
| Db   | 1141 | ATCGCCGTGGCCGTCTGCTGCTGCTGCTGCTCTCTCATCCTCGTTTATTGCCGG             | 1200 |
| Qу   | 1201 | AAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAG        | 1260 |
| Db   | 1201 |  | 1260 |
| Qy , | 1261 | CCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGAC       | 1320 |
| Db   | 1261 |  | 1320 |
| Qу   | 1321 | CTCAGCACCACCACCACCACCAGGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGC          | 1380 |
| Db   | 1321 | CTCAGCACCACCACCACCACCAGGGCAGGCTCTCTGTCCCCGGCAGGATGGGCCCAGC         | 1380 |
| Qу   | 1381 | CCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACA       | 1440 |
| Db   | 1381 |  | 1440 |

l

|   |    |      | ,  |      |
|---|----|------|--|------|
|   | Qy | 1441 | CTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG | 1500 |
|   | Db | 1441 | CTGCACCACGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG   | 1500 |
|   | Qy | 1501 | AACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTC | 1560 |
|   | Db | 1501 |  | 1560 |
|   | Qy | 1561 | CTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGATGCC  | 1620 |
|   | Db | 1561 |  | 1620 |
|   | Qу | 1621 | ATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGG | 1680 |
|   | Db | 1621 |  | 1680 |
|   | Qy | 1681 | TTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGC | 1740 |
|   | Db | 1681 | TTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGC | 1740 |
|   | Qy | 1741 | GTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGAC | 1800 |
|   | Db | 1741 |  | 1800 |
|   | Qу | 1801 | AGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCTGCAC | 1860 |
|   | Db | 1801 |  | 1860 |
|   | Qy | 1861 | CTGGGCGAGGAGGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTAC   | 1920 |
|   | Db | 1861 | CTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTAC | 1920 |
|   | Qy | 1921 | GTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCC | 1980 |
|   | Db | 1921 | GTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCC | 1980 |
|   | Qy | 1981 | GCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC          | 2040 |
|   | Db | 1981 | GCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC          | 2040 |
|   | Qy | 2041 | ATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAG | 2100 |
|   | Db | 2041 | ATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAG | 2100 |
|   | Qy | 2101 | AAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTAC | 2160 |
|   | Db | 2101 | AAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTAC | 2160 |
|   | Qy | 2161 | CACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTT | 2220 |
| ٠ | Db | 2161 |  | 2220 |
|   | Qy | 2221 | GTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCAC | 2280 |
|   | Db | 2221 | GTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCAC | 2280 |

```
2281 TGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGG 2340
Qy
          2281 TGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGG 2340
Db
      2341 GTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGAC 2400
Qy
          2341 GTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGAC 2400
Db
      2401 ACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGC 2460
Qy
          2401 ACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGC 2460
Db
      2461 CCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCA 2520
Qy
          2461 CCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCA 2520
Db
      2521 CCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCAT 2580
Qy
          2521 CCCTGTAGGCGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCAT 2580
Db
      2581 CTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGGGAGGCG 2640
Qу
          2581 CTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGGGAGGCG 2640
Db
      2641 CGGCACTTCCCCAACGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAG 2700
Qy
          2641 CGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAG 2700
Db
      Qу
          Db
RESULT 2
AX451652
                                                PAT 03-JUL-2002
LOCUS
        AX451652
                           2697 bp
                                   DNA
                                         linear
        Sequence 1 from Patent WO0233080.
DEFINITION
        AX451652
ACCESSION
VERSION
        AX451652.1 GI:21698587
KEYWORDS
SOURCE
        Homo sapiens (human)
        Homo sapiens
 ORGANISM
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
 AUTHORS
         Koehler, R.H.
         Regulation of human netrin binding membrane receptor unc5h-1
 TITLE
 JOURNAL
         Patent: WO 0233080-A 1 25-APR-2002;
         Bayer Aktiengesellschaft (DE)
FEATURES
               Location/Qualifiers
               1. .2697
   source
               /organism="Homo sapiens"
                /mol type="unassigned DNA"
                /db xref="taxon:9606"
ORIGIN
                   97.7%; Score 2687.4;
                                    DB 6;
                                          Length 2697;
 Query Match
```

|          |      | imilarity 99.8%;<br>; Conservative |       | . No. 0<br>ismatch |                   | ;         | Indels         | 0;    | Gaps   | 0;  |
|----------|------|------------------------------------|-------|--------------------|-------------------|-----------|----------------|-------|--------|-----|
| Qy<br>Db |      | ATGGCCGTCCGGCCCGGCCC               |       | 111111             | 1111111           | $\square$ | 11111111111    | 1111  | 11111  |     |
| Qу       |      | CGCGGCTCGGGTGCCCAGC                |       | HHHHH              | $\Pi\Pi\Pi\Pi\Pi$ | $\Pi\Pi$  |                | +111  |        |     |
| Db       |      | CGCGGCTCGGGTGCCCAGC                |       |                    |                   |           | •              |       | -      |     |
| Qу       | 166  | GACCTGCTTCCCCACTTCC                |       |                    |                   |           |                |       |        | 225 |
| Db       | 121  | GACCTGCTTCCCCACTTCC'               | rggtg | GAGCCCG            | AGGATGT           | GTA       | CATCGTCAAG     | SAACA | AAGCCA | 180 |
| Qу       | 226  | GTGCTGCTTGTGTGCAAGG                |       |                    |                   |           |                |       |        | 285 |
| Db       | 181  | GTGCTGCTTGTGCAAGG                  |       |                    |                   |           |                |       |        | 240 |
| Qу       | 286  | TGGGTGCGCCAGGTGGACC                |       |                    |                   |           |                |       |        | 345 |
| Db       | 241  |                                    |       |                    |                   |           |                |       |        | 300 |
| Qу       | 346  | ACCATGGAGGTCCGCATTA                |       |                    |                   |           |                |       |        | 405 |
| Db       | 301  |                                    |       |                    |                   |           |                |       |        | 360 |
| Qу       | 406  | GAATACTGGTGCCAGTGCG                |       |                    |                   |           |                |       |        | 465 |
| Db       | 361  |                                    |       |                    |                   |           |                |       |        | 420 |
| Qу       | 466  | TACATCCGCATAGCCAGAT                |       |                    |                   |           |                |       |        | 525 |
| Db       | 421  |                                    |       |                    |                   |           | <br> COGETGGCO |       |        | 480 |
| Qу       | 526  | TCCCTGGAGCAGGCATCG                 |       |                    |                   |           |                |       |        | 585 |
| Db       | 481  |                                    |       |                    |                   |           |                |       |        | 540 |
| Qу       | 586  | GTGGAGTGGCTCCGGAACG.               |       |                    |                   |           |                |       |        | 645 |
| Db       | ·541 |                                    | AGGAC | CTGGTGG            | ACCCGTC           | CCT       | GGACCCCAAI     | GTAT  | ACATC  | 600 |
| Qу       | 646  | ACGCGGGAGCACAGCCTGG                |       |                    |                   |           |                |       |        | 705 |
| Db       | 601  | ACGCGGGAGCACAGCCTGG                |       |                    |                   |           |                |       |        | 660 |
| Qу       | 706  | TGCGTGGCCAAGAACATCG                |       |                    |                   |           |                |       |        | 765 |
| Db       | 661  |                                    |       |                    |                   |           |                |       |        | 720 |
| Qу       | 766  | GTGAACGGTGGGTGGTCGA                |       |                    |                   |           |                |       |        | 825 |
| Db       | 721  | GTGAACGGTGGGTGGTCGA                |       |                    |                   |           |                |       |        | 780 |
| Qу       | 826  | GGCTGGCAGAAACGGAGCC                |       |                    |                   |           |                |       |        | 885 |

| Db | 781  | GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC   | 840  |
|----|------|--|------|
| Qy | 886  | TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC   | 945  |
| Db | 841  | TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTGGACGGCAGC   | 900  |
| Qy | 946  | TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT   | 1005 |
| Db | 901  | TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT   | 960  |
| Qy | 1006 | GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGGG                        | 1065 |
| Db | 961  | GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGGAGTGCCAGGGCACTGACCTGGAC | 1020 |
| Qy | 1066 | ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC   | 1125 |
| Db | 1021 | ACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTGGCC   | 1080 |
| Qy | 1126 | CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATC   | 1185 |
| Db | 1081 | CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCTGCTGCTGCTGCTGCTCATC          | 1140 |
| Qу | 1186 | CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC   | 1245 |
| Db | 1141 | CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC   | 1200 |
| Qу | 1246 | ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC   | 1305 |
| Db | 1201 | ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC   | 1260 |
| Qу | 1306 | ACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCCCGG       | 1365 |
| Db | 1261 | ACCATCCAGCCGGACCTCAGCACCACCACCACCAGGGCAGTCTCTGTCCCCGG          | 1320 |
| Qу | 1366 | CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT   | 1425 |
| Db | 1321 | CAGGATGGGCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT         | 1380 |
| Qу | 1426 | GGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC      | 1485 |
| Db | 1381 | GGCGGCCGCCACACTGCACCACACCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCC      | 1440 |
| Qy | 1486 | CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT   | 1545 |
| Db | 1441 | CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT   | 1500 |
| Qy | 1546 | GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC   | 1605 |
| Db | 1501 | GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTC   | 1560 |
| Qу | 1606 | ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG    | 1665 |
| Db | 1561 | ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG   | 1620 |
| Qу | 1666 | CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC   | 1725 |
| Db | 1621 | CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC   | 1680 |

|   | Qy   | 1726 | TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG   | 1785 |
|---|------|------|---|------|
|   | Db   | 1681 | TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG   | 1740 |
| • | Qy   | 1786 | GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG  | 1845 |
|   | Db . | 1741 |   | 1800 |
|   | Qy   | 1846 | GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG  | 1905 |
|   | Db   | 1801 |   | 1860 |
|   | Qу   | 1906 | GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC  | 1965 |
|   | Db   | 1861 |   | 1920 |
|   | Qy   | 1966 | CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC   | 2025 |
|   | Db   | 1921 | CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC   | 1980 |
|   | Qy   | 2026 | TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG  | 2085 |
|   | Db . | 1981 | TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG  | 2040 |
|   | Qу   | 2086 | GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC  | 2145 |
|   | Db   | 2041 | GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC | 2100 |
|   | Qу   | 2146 | TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG  | 2205 |
|   | Db   | 2101 | TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG  | 2160 |
|   | Qy   | 2206 | AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG  | 2265 |
|   | Db   | 2161 | AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG  | 2220 |
|   | Qу   | 2266 | CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG  | 2325 |
|   | Db   | 2221 | CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG  | 2280 |
|   | Qy   | 2326 | GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC  | 2385 |
|   | Db . | 2281 | GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC  | 2340 |
|   | Qy   | 2386 | AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC  | 2445 |
|   | Db   | 2341 | AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC  | 2400 |
|   | Qy   | 2446 | CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT  | 2505 |
|   | Db   | 2401 | CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT  | 2460 |
|   | Qy   | 2506 | TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC  | 2565 |
|   | Db   | 2461 | TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC  | 2520 |

| QŸ  |  |
|---|--|
| Db  | 2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580   |
| Qy  | 2626 AACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC   |
| Db  | 2581 AACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC   |
| Qy  | 2686 GCTGGACTGGGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2742  |
| Db  | 2641 GCTGGACTGGGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697  |
| RESULT 3 AX527916 LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANIS REFERENCH AUTHORS TITLE JOURNAL FEATURES SOUR ORIGIN | AX527916.1 GI:25172359 . Homo sapiens (human)  Memo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  Herrmann, J.L., Rastelli, L. and Shimkets, R.A. Novel proteins and nucleic acids encoding same and antibodies directed against these proteins Patent: WO 0229038-A 1 11-APR-2002; Curagen Corporation (US) Location/Qualifiers |
|   | Match 97.3%; Score 2676.4; DB 6; Length 2881; ocal Similarity 98.9%; Pred. No. 0; s 2728; Conservative 0; Mismatches 21; Indels 9; Gaps 3  |
| Qу  | 1 CCGCGGGCCCCGCCCGCCCGCCCGCCCGCCCATGCCCGTCCGGCCC 60  |
| Db  |  |
| Qy  | 61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120  |
| Db  | 102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161   |
| Qу  | 121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180   |
| Db  | 162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221   |
| Qу  | 181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240   |
| Db  | 222 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 281   |

\*

| Qy | 241  | AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG | 300  |
|----|------|--|------|
| Db | 282  | AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG | 341  |
| Qу | 301  | GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC | 360  |
| Db | 342  | GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGC | 401  |
| Qу | 361  | ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG | 420  |
| Db | 402  | ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG | 461  |
| Qy | 421  | TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC | 480  |
| Db | 462  | TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC | 521  |
| Qy | 481  | AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC | 540  |
| Db | 522  | AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC | 581  |
| Qу | 541  | ATCGTGCTGCCCTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG      | 600  |
| Db | 582  | ATCGTGCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG      | 641  |
| Qу | 601  | AACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC     | 660  |
| Db | 642  | AACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC     | 701  |
| Qy | 661  | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC | 720  |
| Db | 702  | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC | 761  |
| Qy | 721  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG     | 780  |
| Db | 762  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG     | 821  |
| Qy | 781  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG | 840  |
| Db | 822  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG | 881  |
| Qy | 841  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT | 900  |
| Db | 882  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT | 941  |
| Qy | 901  | GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG    | 957  |
| Db | 942  | GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG | 1001 |
| Qy | 958  | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC | 1017 |
| Db | 1002 | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC | 1061 |
| Qу | 1018 | CCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT   | 1077 |
| Db | 1062 | CCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT  | 1121 |
| Ov | 1078 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC | 1137 |

.

| Db   | 1122 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC | 1181 |
|------|------|--|------|
| QУ   | 1138 | CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGC | 1197 |
| Db   | 1182 | CTCATCGCCGTGGCCTGCTGCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGC        | 1241 |
| Qу   | 1198 | CGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC  | 1257 |
| Db   | 1242 | CGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC  | 1301 |
| QУ   | 1258 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG | 1317 |
| Db   | 1302 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG | 1361 |
| Qу   | 1318 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC    | 1377 |
| Db   | 1362 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC    | 1418 |
| QУ   | 1378 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC | 1437 |
| Db   | 1419 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC | 1478 |
| Qу   | 1438 | ACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC | 1497 |
| Db   | 1479 | ACACTGCACCACACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC          | 1538 |
| Qу   | 1498 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC | 1557 |
| Db   | 1539 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC | 1598 |
| Qу   | 1558 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT  | 1617 |
| Db   | 1599 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT  | 1658 |
| Qу   | 1618 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG | 1677 |
| Db   | 1659 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG | 1718 |
| Qу   | 1678 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT | 1737 |
| Db   | 1719 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT | 1778 |
| Qy   | 1738 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1797 |
| Db   | 1779 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1838 |
| Qу   | 1798 | GACAGCTGGAGCCTCCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTG         | 1854 |
| Db . | 1839 | GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG | 1898 |
| Qу   | 1855 | CTGCACCTGGGCGAGGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC   | 1914 |
| Db   | 1899 | CTGCACCTGGGCGAGGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC   | 1958 |
| Qу   | 1915 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG | 1974 |

·

.

| Db       | 1959 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG   | 2018 |
|----------|------|--|------|
| Qу       | 1975 | GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC  | 2034 |
| Db       | 2019 | GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC  | 2078 |
| Qу       | 2035 | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG   | 2094 |
| Db       | 2079 | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG   | 2138 |
| Qу       | 2095 | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC   | 2154 |
| Db       | 2139 | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC   | 2198 |
| Qy.      |      | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG   |      |
| Db       | 2199 | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG   | 2258 |
| Qy       |      | CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC   |      |
| Db       |      | CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC   |      |
| QУ       |      | TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG   |      |
| Db       |      | TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG   |      |
| Qy<br>D' |      | CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC   |      |
| Db       |      | AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG   |      |
| Qy<br>Db |      | AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG  AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG |      |
| Qу       |      | GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG   |      |
| Db       |      |  |      |
| Qy       | 2515 | GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC   | 2574 |
| Db       | 2559 |  | 2618 |
| Qу       | 2575 | AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG   | 2634 |
| Db       | 2619 | AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG   | 2678 |
| Qy       | 2635 | GAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTG   | 2694 |
| Db       | 2679 |  | 2738 |
| Qу       | 2695 | GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCC   | 752  |
| Db       | 2739 |  | 796  |

```
CO730306 '
                             2784 bp
                                             linear
                                                    PAT 03-FEB-2004
LOCUS
                                      DNA
         Sequence 16240 from Patent W002068579.
DEFINITION
         CO730306
ACCESSION
         CQ730306.1 GI:42303801
VERSION
KEYWORDS
SOURCE
         Homo sapiens (human)
 ORGANISM
         Homo sapiens
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
         Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
 AUTHORS
 TITLE
         Kits, such as nucleic acid arrays, comprising a majority of
         humanexons or transcripts, for detecting expression and other uses
         thereof
 JOURNAL
         Patent: WO 02068579-A 16240 06-SEP-2002;
         PE Corporation (NY) (US)
                 Location/Qualifiers
FEATURES
                 1. .2784
    source
                 /organism="Homo sapiens"
                 /mol type="unassigned DNA"
                 /db xref="taxon:9606"
ORIGIN
 Query Match
                     88.5%; Score 2435.4; DB 6;
                                             Length 2784;
 Best Local Similarity
                     93.5%;
                           Pred. No. 0;
 Matches 2646; Conservative
                          0;
                             Mismatches
                                         6;
                                             Indels 177; Gaps
                                                               3:
        46 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 105
Qу
           Db
         1 ATGGCCGTCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTTGGCTC 60
        106 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 165
Qy
           61 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 120
Db
        166 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 225
Qy
           121 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 180
Db
        226 GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 285
Qу
           181 GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 240
Db
        286 TGGGTGCGCCAGGTGACCACGTGATCGAGCGCACAGACGGGAGCAGTGGGCTGCCC 345
Qу
           Db
        241 TGGGTGCCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 300
        346 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 405
Qу
           301 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 360
Db
        406 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 465
Qy
           361 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 420
Db
Qy
        466 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 525
           111111111111
```

| Db | 421  | TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG         | 480  |
|----|------|--|------|
| Qу | 526  | TCCCTGGAGCAGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG          | 585  |
| Db | 481  | TCCCTGGAGCAGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG          | 540  |
| Qу | 586  | GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC         | 645  |
| Db | 541  | GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC         | 600  |
| Qy | 646  | ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC         | 705  |
| Db | 601  | ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC         | 660  |
| Qy | 706  | TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC         | 765  |
| Db | 661  | TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTAC            | 720  |
| QУ | 766  | ·  | 765  |
| Db | 721  | ${\tt GGTGGGCCCCGGGACTCCCTGGTCACAGGGAGAGGCACTGCGGTGCCCCTGGGCAGTGAC}$ | 780  |
| Qу | 766  | GTGAACGGTGGGTCGACGTGGACCGAGTGG                                       | 798  |
| Db | 781  | ATGTGGCTGTCCTCTCTCTCCGGCCAGTGAACGGTGGGTCGACGTGGACCGAGTGG             | 840  |
| Qу | 799  | TCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGCACCAAC                              | 858  |
| Db | 841  | TCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAAC         | 900  |
| Qy | 859  | CCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGC         | 918  |
| Db | 901  | CCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGC         | 960  |
| Qy | 919  | GCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGG         | 978  |
| Db | 961  | GCCACCCTGTGCCCAGTGGACGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGG          | 1020 |
| Qy | 979  | CTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGG         | 1038 |
| Db | 1021 | CTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGG         | 1080 |
| Qy | 1039 | GAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACAC         | 1098 |
| Db | 1081 | GAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACAC         | 1140 |
| Qу | 1099 | AGTGCTTCTGGCCCT  | 1113 |
| Db | 1141 | AACTCCTACACCCCTGCCCCCACCAAGGCCATGCTGTCTCCCGCAGCTGCTTCTGGCCCT         | 1200 |
| Qу | 1114 | GAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC                    | 1173 |
| Db | 1201 | GAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC                    | 1260 |
| Qу | 1174 | CTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGAC         | 1233 |
| Db | 1261 | CTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGAC         | 1320 |

| Qy | 1234 | TCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAAC      | 1293 |
|----|------|---|------|
| Db | 1321 | TCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAAC      | 1380 |
| Qу | 1294 | CCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCACCACCAGGGCAGT | 1353 |
| Db | 1381 | CCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCAGGGCAGT       | 1440 |
| Qy | 1354 | CTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTC      | 1413 |
| Db | 1441 | CTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTC      | 1500 |
| Qу | 1414 | AGCCCCTGGGTGGCGGCCGCCACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAG         | 1473 |
| Db | 1501 | AGCCCCTGGGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAG          | 1560 |
| Qу | 1474 | GAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGC      | 1533 |
| Db | 1561 | GAGTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGC      | 1620 |
| Qу | 1534 | AACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGT      | 1593 |
| Db | 1621 | AACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGA      | 1680 |
| Qy | 1594 | ATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTC       | 1653 |
| Db | 1681 | ATCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTC       | 1740 |
| Qy | 1654 | ACGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTG                      | 1713 |
| Db | 1741 | ACGCTGCACAAGCCGGAAGACGTG  | 1764 |
| Qy | 1714 | CCCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATG      | 1773 |
| Db | 1765 | AGCTGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATG                | 1815 |
| Qy | 1774 | GACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGC      | 1833 |
| Db | 1816 | GACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGC      | 1875 |
| Qу | 1834 | GAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTAC      | 1893 |
| Db | 1876 | GAGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTAC      | 1935 |
| Qy | 1894 | TGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTG      | 1953 |
| Db | 1936 | TGCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTG      | 1995 |
| Qy | 1954 | GTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCG      | 2013 |
| Db | 1996 | GTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCCCCTCAAGCTGCTTCTGTTTGCGCCG      | 2055 |
| Qy | 2014 | GTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC                    | 2073 |
| Db | 2056 | GTGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC                    | 2115 |

| Qу       | 2074 | GCACTCAAGGAGGTGCTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCA | 2133 |
|----------|------|---|------|
| Db .     | 2116 | GCACTCAAGGAGGTGCTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCA | 2175 |
| Qy       | 2134 | CGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC             | 2193 |
| Db       | 2176 | CGGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC             | 2235 |
| Qy       | 2194 | AGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATC  | 2253 |
| Db       | 2236 | AGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATC  | 2295 |
| Qу       | 2254 | TGGAATGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGC    | 2313 |
| Db       | 2296 | TGGAATGGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGC   | 2355 |
| Qy       | 2314 | ACTAGTGACCTGCCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTC      | 2373 |
| Db       | 2356 | ACTAGTGACCTGCCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTC      | 2415 |
| Qy       | 2374 | AGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGT  | 2433 |
| Db       | 2416 | AGCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGT  | 2475 |
| Qy       | 2434 | GAAGCGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGG   | 2493 |
| Db       | 2476 | GAAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGG  | 2535 |
| Qу       | 2494 | CAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTG  | 2553 |
| Db       | 2536 | CAGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTG  | 2595 |
| Qy       | 2554 | GCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACA   | 2613 |
| Db       | 2596 | GCCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACA  | 2655 |
| Qy,      | 2614 | GCCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGC    | 2673 |
| Db       | 2656 | GCCATGATCCTCAACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC     | 2715 |
| Qy       | 2674 | GCTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCT  | 2733 |
| Db       | 2716 | GCTGCAGCAGTGGCCTGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCT       | 2775 |
| Qу       | 2734 | GAGTGCTGA 2742  |      |
| Db       | 2776 | GAGTGCTGA 2784  |      |
| RESULT 5 |      |   |      |

LOCUS AX367094 3580 bp DNA linear PAT 16-FEB-2002

DEFINITION Sequence 13 from Patent W00198354.

ACCESSION AX367094

VERSION AX367094.1 GI:18855296

KEYWORDS

SOURCE Homo sapiens (human)

```
ORGANISM
         Homo sapiens
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
         Griffin, J.A., Kallick, D.A., Tribouley, C.M., Yue, H., Nguyen, D.B.,
 AUTHORS
         Tang, Y.T., Lal, P., Policky, J.L., Azimzai, Y., Lu, D.A., Graul, R.,
         Yao, M.G., Burford, N., Hafalia, A.J., Baughn, M.R., Bandman, O.,
         Patterson, C., Yang, J., Xu, Y., Warren, B.A., Ding, L. and
         Sanjanwala, M.S.
         Receptors
 TTTLE.
         Patent: WO 0198354-A 13 27-DEC-2001;
 JOURNAL
         Incyte Genomics, Inc. (US)
                 Location/Qualifiers
FEATURES
                 1. .3580
    source
                 /organism="Homo sapiens"
                 /mol type="unassigned DNA"
                 /db xref="taxon:9606"
                 /note="Incyte ID No: 6052371CB1"
ORIGIN
                     85.6%;
                            Score 2356; DB 6;
                                            Length 3580;
 Query Match
 Best Local Similarity
                            Pred. No. 0;
                     93.6%;
 Matches 2537; Conservative
                           0: Mismatches
                                             Indels
                                                    168;
         43 GCCATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGG 102
Qу
           1 GCCATGGCCGTCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGG 60
Db
        103 CTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAAC 162
Qу
           61 CTCCGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAAC 120
Db
        163 CCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAG 222
Qу
           121 CCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAG 180
Db
        223 CCAGTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGG 282
Qy
           181 CCAGTGCTGCTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGG 240
Db
        283 GAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCACACAGACGGGAGCAGTGGGCTG 342
Qy
           241 GAGTGGGTGCCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTG 300
Db
        343 CCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTG 402
Qу
           301 CCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTG 360
Db
        403 GAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAG 462
Qу
           361 GAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAG 420
Db
        463 GCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAG 522
Qу
                            421 GCCTACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAG 480
Db
        523 GTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCC 582
Qу
```

| Db | 481  | GTGTCCCTGGAGCAGGGCATCGTGCCTGCCGTCCACCGGAGGGCATCCCTCCAGCC           | 540  |
|----|------|--|------|
| Qу | 583  | GAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATAC       | 642  |
| Db | 541  |  | 600  |
| Qу | 643  | $\tt ATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTAC$ | 702  |
| Db | 601  |  | 660  |
| Qу | 703  | ACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTCATCGTC         | 762  |
| Db | 661  |  | 720  |
| Qу | 763  | TACGTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGG           | 822  |
| Db | 721  | <br>TACGTG   | 726  |
| Qу | 823  | CGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCT       | 882  |
| Db | 727  |  | 726  |
| Qу | 883  | TTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGC       | 942  |
| Db | 727  | GACGGC   | 732  |
| Qу | 943  | AGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC       | 1002 |
| Db | 733  | AGCTGGAGCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGC        | 792  |
| Qу | 1003 | CGTGAGTGCTCTGACCCAGCACCCGCAACGGAGGGGAGG                            | 1062 |
| Db | 793  | CGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGTGCCAGGGCACTGACCTG          | .852 |
| Qу | 1063 | GACACCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTG        | 1122 |
| Db | 853  | GACACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTG       | 912  |
| Qу | 1123 | GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTC       | 1182 |
| Db | 913  | GCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGC                           | 972  |
| Qу | 1183 | ATCCTCGTTTATTGCCGGAAGAAGGAGGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT      | 1242 |
| Db | 973  | ATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATT       | 1032 |
| Qy | 1243 | CTCACCTCAGGCTTCCAGCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG        | 1302 |
| Db | 1033 | CTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTG       | 1092 |
| Qy | 1303 | CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCAGGGCAGTCTCTGTCCC        | 1362 |
| Db | 1093 | CTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCC           | 1152 |
| Qу | 1363 | CGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG       | 1422 |

|   | Db   | 1153 | CGGCAGGATGGGCCCAAGCTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTG        | 1212 |
|---|------|------|---|------|
|   | Qy   | 1423 | GGTGGCGGCCACACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTC      | 1482 |
|   | Db   | 1213 | GGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTC     | 1272 |
|   | Qy   | 1483 | TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC  | 1542 |
|   | Db   | 1273 | TCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACC  | 1332 |
| ٠ | Qу   | 1543 | TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTC  | 1602 |
|   | Db   | 1333 | TATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTC  | 1392 |
|   | Qy   | 1603 | CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC   | 1662 |
|   | Db   | 1393 | CTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCAC   | 1452 |
|   | Qy   | 1663 | AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT  | 1722 |
|   | Dp , | 1453 | AAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTT  | 1512 |
|   | Qy   | 1723 | AGCTGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT   | 1782 |
|   | Db   | 1513 | AGCTGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGT      | 1572 |
|   | Qy   | 1783 | GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC  | 1842 |
|   | Db   | 1573 | GGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGC  | 1632 |
|   | Qy   | 1843 | TGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTG  | 1902 |
|   | Db   | 1633 | TGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTG  | 1692 |
|   | Qy   | 1903 | GAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAG  | 1962 |
| • | Db   | 1693 | GAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAG  | 1752 |
|   | Qy _ | 1963 | GCCCTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGC   | 2022 |
|   | Db   | 1753 | GCCCTCAGCGTGGCTGCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGC    | 1812 |
|   | Qу   | 2023 | ACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAG  | 2082 |
|   | Db   | 1813 | ACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAG  | 1872 |
|   | Qу   | 2083 | GAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTG | 2142 |
|   | Db   | 1873 | GAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTG | 1932 |
|   | Qу   | 2143 | CACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC                      | 2202 |
|   | Db   | 1933 | CACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC                      | 1992 |
|   | Qy   | 2203 | TGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGC  | 2262 |
|   | Db   | 1993 | TGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGC  | 2052 |

| Qу                | 2263 | ACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGAC 2322                |
|-------------------|------|--|
| Db                | 2053 | ACGCAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGAC 2112                |
| Qу                | 2323 | CTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAAC 2382                |
| Db                | 2113 |  |
| QУ                | 2383 | TTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGGCTCTGGAGAGTGAAGCGGGG 2442                   |
| Db                | 2173 | TTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGG 2232                |
| Qy                | 2443 | GTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATA 2502                |
| Db                | 2233 |  |
| Qу                | 2503 | ATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAA 2562                |
| Db                | 2293 | ATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAA 2352                |
| Qy                | 2563 | CTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATC 2622                 |
| Db                | 2353 | CTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATC 2412                |
| Qy                | 2623 | CTCAACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC                                 |
| Db                | 2413 | CTCAACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC                                 |
| Qу                | 2683 | GTGGCTGGACTGGGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2742                    |
| Db .              | 2473 |  |
| Qy                | 2743 | GGCCGGCCAG 2752  |
| Dp.               | 2533 |  |
| •                 |      |  |
| RESULT 6          | 2    |  |
| LOCUS             | M    | MU487852 3992 bp mRNA linear ROD 24-SEP-2002                                     |
| DEFINITI          |      | (us musculus mRNA for netrin receptor Unc5h1 (Unc5h1 gene).                      |
| ACCESSION VERSION |      | J487852<br>J487852.1 GI:22035783   |
| KEYWORDS          |      | etrin receptor Unc5h1; Unc5h1 gene.  |
| SOURCE            |      | (us musculus (house mouse)   |
| ORGANI            |      | ius musculus   |
|                   |      | ukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;                 |
| REFERENC          |      | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.               |
| AUTHOR            |      | Engelkamp, D.  |
| TITLE             | C    | Cloning of three mouse Unc5 genes and their expression patterns at uid-gestation |
| JOURNA            |      | lech. Dev. 118 (1-2), 191-197 (2002)   |
| MEDLIN            |      | 2239710  |
| PUBME             | D 1  | 2351186  |

REFERENCE 2 (bases 1 to 3992)

AUTHORS Engelkamp, D.

TITLE Direct Submission

JOURNAL Submitted (15-MAY-2002) Neuroanatomy, Max Planck Institute for

Brain Research, Deutschordenstrasse 46, Frankfurt 60528, GERMANY

FEATURES Location/Qualifiers

source 1. .3992

/organism="Mus musculus"

/mol type="mRNA"

/db xref="taxon:10090"

gene 1. .3992

/gene="Unc5h1"

CDS 232..2928 /gene="Unc5h1

/gene="Unc5h1"
/codon start=1

/product="netrin receptor Unc5h1"

/protein\_id="CAD32250.1"
/db\_xref="GI:22035784"
/db\_xref="GOA:Q8K1S4"

/db xref="UniProt/TrEMBL:Q8K1S4"

/translation="MAVRPGLWPALLGIVLTAWLRGSGAQQSATVANPVPGANPDLLP HFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTM EVRINVSRQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEV SLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTAN YTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLN GGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEEC RGADLDTRNCTSDLCLHTSSGPEDVALYIGLVAVAVCLILLLLVLVLIYCRKKEGLDS DVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKFQL SNGHLLSPLGSGRHTLHHSSPTSEAEDFVSRLSTQNYFRSLPRGTSNMAYGTFNFLGG RLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGV LLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEESPSHLYYCQLEAGAC YVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKEVV QLEKQLGGQL1QEPRVLHFKDSYHNLRLS1HDVPSSLWKSKLLVSYQE1PFYH1WNGT QQYLHCTFTLERVNASTSDLACKVWVWQVEGDGQSFNINFNITKDTRFAEMLALESEG GVPALVGPSAFKIPFLIRQKIITSLDPPCSRGADWRTLAQKLHLDSHLSFFASKPSPT AMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC"

#### ORIGIN

Query Match 84.4%; Score 2323.2; DB 10; Length 3992; Pred. No. 0; Best Local Similarity 90.3%; 0; Mismatches 0; Matches 2484; Conservative 0; 268; Indels Gaps 1 CCGCGGGGCCCGCCCGCCCGCCCGCCCGCCGCGCCATGCCCGTCCGGCCC 60 Qу 187 CCGCCGCGGGGCCCCTAGTCCAGCCCGCTCGTCCGCCGCGGCCATGGCCGTCCGGCCC 246 Db

Qy 61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120

Db 247 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCACTGCCTTGGTTCGGGTTCGGGTGCC 306

Qy 121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTĞCCAACCCGGACCTGCTTCCCCAC 180

Db 307 CAGCAGAGTGCCACAGTGGCCAACCCAGTGCCTGGTGCCAACCCAGACCTGCTGCCCCAC 366

Qy 181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240

Db 367 TTCCTGGTAGAGCCGGAGGACGTGTACATTGTCAAGAACAAGCCCGTGCTGCTGTGTGC 426

| QΆ | 241  | AAGGCCGTGCCCGCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG  | 300  |
|----|------|--|------|
| Db | 427  | AAGGCTGTGCCCGCCACCCAGATCTTCTTCAAGTGCAACGGGGAATGGGTTCGCCAGGTC | 486  |
| Qу | 301  | GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC | 360  |
| Db | 487  | GATCACGTCATCGAACGCAGCAGCAGCAGCAGCAGCGGATTGCCAACCATGGAGGTCCGG | 546  |
| Qу | 361  | ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG | 420  |
| Db | 547  | ATCAACGTATCAAGGCAGCAGGTCGAGAAAGTGTTTGGGCTGGAGGAGTACTGGTGCCAG | 606  |
| Qу | 421  | TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC | 480  |
| Db | 607  | TGTGTGGCATGGAGCTCCTCAGGAACCACCAAAAGCCAGAAGGCCTACATCCGGATTGCC | 666  |
| Qу | 481  | AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC | 540  |
| Db | 667  | TATTTGCGCAAGAACTTTGAGCAGGAGCCGCTGGCCAAGGAAGTGTCACTGGAGCAAGGC | 726  |
| Qу | 541  | ATCGTGCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG      | 600  |
| Db | 727  | ATTGTGCTACCTTGTCGCCCCCGGAAGGAATCCCCCCAGCTGAGGTGGAGTGGCTCCGA  | 786  |
| Qу | 601  | AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC | 660  |
| Db | 787  | AATGAGGACCTCGTGGACCCCTCGACCCCAATGTGTACATCACACGGGAGCACAGC     | 846  |
| Qу | 661  | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC | 720  |
| Db | 847  | CTAGTCGTGCGGCAGGCCCGCCTGGCCGACACTACACCTGCGTGGCCAAGAAC        | 906  |
| Qу | 721  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG     | 780  |
| Db | 907  | ATCGTGGCCCGTCGCCGAAGCGCCTCTGCGGCCGTCATTGTTTATGTGAACGGTGGGTG  | 966  |
| Qу | 781  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG | 840  |
| Db | 967  | TCGACGTGGACCGAGTGGTCCGTCTGCAGTGCCAGCTGTGGGCGTGGCTGGC         | 1026 |
| Qу | 841  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT | 900  |
| Db | 1027 | AGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTCTGTGAGGGGCAGAAT | 1086 |
| Qу | 901  | GTCCAGAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGGAGC  | 960  |
| Db | 1087 | GTCCAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGGAGCTGGAGCCCATGGAGT | 1146 |
| Qу | 961  | AAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCA | 1020 |
| Db | 1147 |  | 1206 |
| Qу | 1021 | GCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACC  | 1080 |
| Db | 1207 | GCGCCCGCAACGGAGGTGAGGAGTGCCGGGGTGCTGACCTGGACACCCGCAACTGTACC  | 1266 |
| Ov | 1081 | AGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTC | 1140 |

| Db | 1267 |  | 1326 |
|----|------|--|------|
| Qу | 1141 | ATCGCCGTGGCCGTCTGCTGCTGCTGCTGCTCATCCTCGTTTATTGCCGG           | 1200 |
| Db | 1327 | GTCGCCGTGGCCGTGTGCCTCATCTTGCTGCTGGTCCTCGTCCTCATCTACTGCCGC    | 1386 |
| Qy | 1201 | AAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAG | 1260 |
| Db | 1387 | AAGAAGGAAGGACTGGACTCAGACGTGGCTGACTCATCCTTACCTCAGGCTTCCAG     | 1446 |
| Qy | 1261 | CCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGAC | 1320 |
| Db | 1447 | CCTGTCAGCATCAAGCCCAGCAAAGCAGACAATCCCCATCTGCTCACCATCCAACCGGAC | 1506 |
| Qу | 1321 | CTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGC    | 1380 |
| Db | 1507 | CTCAGCACCACCACCACCTACCAGGGCAGCCTGTGTCCCCGGCAGGATGGACCCAGC    | 1566 |
| Qу | 1381 | CCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACA | 1440 |
| Db | 1567 | CCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTGGGCAGTGGCCGCCATACG | 1626 |
| Qу | 1441 | CTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG | 1500 |
| Db | 1627 | CTGCACCACACCTCTGAGGCTGAGGACTTCGTCTCCCGCCTCTCCACCCAA          | 1686 |
| Qу | 1501 | AACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTC | 1560 |
| Db | 1687 | AACTACTTTCGTTCTCTGCCCCGCGGTACCAGCAACATGGCCTATGGGACCTTCAACTTC | 1746 |
| Qy | 1561 | CTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGATGCC  | 1620 |
| Db | 1747 | CTCGGGGGCCGGCTGATGATCCCTAACACAGGAATCAGCCTCCTCATACCCCCGGACGCC | 1806 |
| Qу | 1621 | ATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGG | 1680 |
| Db | 1807 | ATCCCCCGAGGAAAGATCTACGAGATCTACCTCACTCTGCACAAGCCAGAAGACGTGAGG | 1866 |
| Qy | 1681 | TTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGC | 1740 |
| Db | 1867 | TTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCTATCGTTAGCTGTGGGCCCCCAGGA | 1926 |
| Qy | 1741 | GTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGAC | 1800 |
| Db | 1927 | GTCCTGCTCACCCGGCCAGTCATCCTTGCCATGGACCACTGCGGGGAGCCCAGTCCCGAC | 1986 |
| Qy | 1801 | AGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCTGCAC | 1860 |
| Db | 1987 | AGCTGGAGCCTGCGCCTCAAAAAGCAGTCCTGTGAGGGCAGCTGGGAGGACGTGCTGCAC | 2046 |
| Qy | 1861 | CTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTAC | 1920 |
| Db | 2047 | CTTGGTGAGGAGTCGCCCTCTCATCTCTACTACTGCCAGCTGGAGGCCGGGGCCTGCTAT | 2106 |
| Qу | 1921 | GTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCC | 1980 |

| Db   | 2107 | GTCTTCACCGAGCAGCTAGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCC         | 2166 |
|------|------|--|------|
| Qу   | 1981 | GCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC                  | 2040 |
| Db   | 2167 | ACCAAGCGCCTCAGGCTCCTTCTGTTTGCCCCTGTGGCCTGTACGTCCCTCGAGTACAAC         | 2226 |
| Qу   | 2041 | ATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAG         | 2100 |
| Db   | 2227 | ATCCGAGTGTACTGCCTGCACGACACCCACGATGCTCTCAAGGAGGTGGTGCAGCTGGAG         | 2286 |
| Qу   | 2101 | AAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTAC          | 2160 |
| Db   | 2287 | AAGCAGCTGGGTGGACAGCTGATCCAGGAGCCCCGTGTCCTGCACTTCAAAGACAGTTAC         | 2346 |
| Qу   | 2161 | CACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTT         | 2220 |
| Db   | 2347 | CACAACCTACGTCTGTCCATCCACGACGTGCCCAGCTCCCTGTGGAAGAGCAAGCTCCTT         | 2406 |
| Qу   | 2221 | GTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCAC         | 2280 |
| Db   | 2407 | GTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAATGGCACTCAGCAGTATCTGCAC         | 2466 |
| Qу   | 2281 | TGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGG         | 2340 |
| Db   | 2467 | TGCACCTTCACCCTGGAGCGCGTCAATGCCAGCACCAGCGACCTGGCCTGCAAGGTGTGG         | 2526 |
| Qу   | 2341 | GTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGAC         | 2400 |
| Db   | 2527 | ${\tt GTGTGGCAGGTGGAGGGAGATGGACAGAGCTTCAACATCAACTTTAACATCACTAAGGAC}$ | 2586 |
| Qу   | 2401 | ACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGC         | 2460 |
| Db   | 2587 | ACGAGGTTTGCTGAAATGCTGGCTCTGGAGAGTGAAGGGGGGGTCCCAGCCCTGGTGGGC         | 2646 |
| Qу   | 2461 | CCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCA         | 2520 |
| Db   | 2647 | CCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAGATCATTACCAGCCTGGACCCA         | 2706 |
| Qу   | 2521 | CCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCAT         | 2580 |
| Db   | 2707 | CCCTGCAGCCGGGGCGCCGACTGGCGAACTCTAGCCCAGAAACTTCACCTGGACAGCCAT         | 2766 |
| Qу   | 2581 | CTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGGGAGGCG          | 2640 |
| Db   | 2767 | CTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTCAACCTATGGGAGGCG         | 2826 |
| Qy   | 2641 | CGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAG         | 2700 |
| Db   | 2827 | CGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCCGGACTGGGCCAG                     | 2886 |
| Qy . | 2701 | CCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCC                   |      |
| Db   | 2887 | CCAGATGCTGGCCTCTTCACCGTGTCAGAGGCCGAGTGCTGAGACCAGCCAG                 |      |

```
3014 bp
                                        DNA
                                                       PAT 27-AUG-2002
LOCUS
          BD057524
                                               linear
DEFINITION
          Netrin receptors.
ACCESSION
          BD057524
          BD057524.1 GI:22603130
VERSION
KEYWORDS
          JP 2001505062-A/1.
SOURCE
          synthetic construct
 ORGANISM
          synthetic construct
          other sequences; artificial sequences.
             (bases 1 to 3014)
REFERENCE
          Lavigne, M.T., Leonardo, D.E., Hinck, L., Masu, M. and Masu, K.K.
 AUTHORS
          Netrin receptors
 TITLE
 JOURNAL
          Patent: JP 2001505062-A 1 17-APR-2001;
          THE REGENTS OF THE UNIV OF CALIFORNIA
          PN
              JP 2001505062-A/1
COMMENT
          PD
              17-APR-2001
              19-FEB-1998 JP 1998536840
          PF
          PR
              19-FEB-1997 US
                              08/808982
              MARC TESSIER LAVIGNE, DAVID E LEONARDO, LINDSAY HINCK, MASAYUKI
          PΙ
          PΙ
          PΙ
              KAZUKO KEINO MASU
              C07K1/00,C07K14/00,C07K17/00,C07H21/02,C07H21/04,G01N33/53 CC
          PC
           Strandedness: Double;
          CC
              Topology: Linear;
          FΗ
              Key
                            Location/Qualifiers.
FEATURES
                  Location/Qualifiers
                  1. .3014
    source
                  /organism="synthetic construct"
                  /mol type="genomic DNA"
                  /db xref="taxon:32630"
ORIGIN
                      82.1%;
                             Score 2259; DB 6; Length 3014;
 Query Match
                      89.7%; Pred. No. 0;
 Best Local Similarity
 Matches 2427; Conservative
                            0; Mismatches 280; Indels
         46 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 105
Qy
            1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60
Db
        106 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 165
Qу
            61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120
Db
        166 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 225
Qy
            121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180
Db
        226 GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 285
Qу
            181 GTGTTGTTGGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240
Db
        286 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 345
Qу
            -1111
        241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCCGACAGCAGCAGCGGATTGCCA 300
Db
        346 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 405
Qy
```

|   | Db   | 301  | ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG       | 360  |
|---|------|------|--|------|
|   | Qy . | 406  | GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC       | 465  |
|   | Db . | 361  | GAATACTGGTGCCAGTGTGGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC        | 420  |
|   | QУ   | 466  | TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG       | 525  |
|   | Db   | 421  | TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG       | 480  |
|   | Qу   | 526  | TCCCTGGAGCAGGCATCGTGCTGCCCTGCCGTCCACCGGAGGCATCCCTCCAGCCGAG         | 585  |
|   | Db   | 481  | TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCAGAAGGAATCCCCCCAGCTGAG        | 540  |
|   | Qy . | 586  | GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC       | 645  |
|   | Db   | 541  | GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCGATCCCAATGTGTACATC          | 600  |
|   | Qу   | 646  | ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC       | 705  |
|   | Db   | 601  | ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC       | 660  |
|   | Qу   | 706  | TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC       | 765  |
|   | Db   | 661  | TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT       | 720  |
|   | Qу   | 766  | GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC           | 825  |
|   | Db   | 721  | GTGAACGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT           | 780  |
|   | Qу   | 826  | GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC       | 885  |
|   | Db   | 781  | GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC       | 840  |
|   | Ωу   | 886  | TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC       | 945  |
|   | Db   | 841  | TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGGAGC       | 900  |
|   | Qy . | 946  | TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT       | 1005 |
| • | Db . | 901  | $\tt TGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGCCGC$ | 960  |
|   | Qу   | 1006 | GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGGG                            | 1065 |
|   | Db   | 961  | GAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTGGAC       | 1020 |
|   | Qy   | 1066 | ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC       | 1125 |
|   | Db   | 1021 | ACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTGGCT       | 1080 |
|   |      |      | CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATC       |      |
|   | Db   | 1081 | CTCTACATCGGCCTTGTCGCTGTGGCTGTGCCTCTTCTTGCTGTTGCTGGCCCTTTGGA        | 1140 |
|   | Qy   | 1186 | CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC       | 1245 |
|   | Db   | 1141 | CTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATCCTC       | 1200 |

.

| Qу | 1246 | ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC | 1305 |
|----|------|--|------|
| Db | 1201 | ACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTGCTC | 1260 |
| Qу | 1306 | ACCATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCCCGG     | 1365 |
| Db | 1261 | ACCATCCAGCCAGACCTCAGCACCACCACCACCACCAGGGCAGTCTATGTTCGAGG     | 1320 |
| Qу | 1366 | CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT | 1425 |
| Db | 1321 | CAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTGGGG | 1380 |
| Qу | 1426 | GGCGGCCGCCACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC   | 1485 |
| Db | 1381 | AGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTCTCC | 1440 |
| Qу | 1486 | CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT | 1545 |
| Db | 1441 | CGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCCTAC | 1500 |
| Qу | 1546 | GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC | 1605 |
| Db | 1501 | GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTCCTC | 1560 |
| Qу | 1606 | ATCCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG | 1665 |
| Db | 1561 | ATACCCCGGATGCCATCCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCACAAG  | 1620 |
| Qу | 1666 | CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC | 1725 |
| Db | 1621 | CCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTTAGC | 1680 |
| Qy | 1726 | TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG  | 1785 |
| Db | 1681 | TGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGTGGA | 1740 |
| Qy | 1786 | GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG | 1845 |
| Db | 1741 | GAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGTTGG | 1800 |
| Qy | 1846 | GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG | 1905 |
| Db | 1801 | GAGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAGCTGGAG | 1860 |
| Qy | 1906 | GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC | 1965 |
| Db | 1861 | GCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGAGAGGCC | 1920 |
| Qy | 1966 | CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC  | 2025 |
| Db | 1921 | CTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCCTGTACG | 1980 |
| Qy | 2026 | TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG | 2085 |
| Db | 1981 | TCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTCAAGGAG | 2040 |

| Qу   | 2086 | GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC 21 | 45 |
|------|------|--|----|
| Db   | 2041 | GTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTCCTGCAC 21  | 00 |
| Qу   | 2146 | TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 22  | 05 |
| Db   | 2101 | TTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCCCTGTGG 21  | 60 |
| Qу   | 2206 | AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 22  | 65 |
| Db   | 2161 | AAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTTACCACATCTGGAACGGCACC 22 | 20 |
| Qу   | 2266 | CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 23  | 25 |
| Db   | 2221 | CAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGCGACCTG 22  | 80 |
| Qy   | 2326 | GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 23  | 85 |
| Db . | 2281 | GCCTGCAAGGTGTGGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATCAACTTC 23       | 40 |
| Qy   | 2386 | AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 24  | 45 |
| Db   | 2341 | AACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGGGGGGTC 24  | 00 |
| Qy   | 2446 | CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 25  | 05 |
| Db   | 2401 | CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAGATCATC 24  | 60 |
| Qy   | 2506 | TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 25  | 65 |
| Db   | 2461 | GCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAGAAACTT 25  | 20 |
| Qу   | 2566 | CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 26  | 25 |
| Db   | 2521 | CACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTC 25  | 80 |
| Qy   | 2626 | AACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC                    | 85 |
| Db   | 2581 | AACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCAGCTGTG 26  | 40 |
| Qy . | 2686 | GCTGGACTGGGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGC 27      | 45 |
| Db   | 2641 | GCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGTTGAGAC 27  | 00 |
| Qy   | 2746 | CGGCCAG 2752   |    |
| Db   | 2701 | CAGCCAG 2707   |    |
|      |      |  |    |

# RESULT 8 AX268596

LOCUS AX268596 2697 bp DNA linear PAT 29-OCT-2001

DEFINITION Sequence 15 from Patent WO0175440.

ACCESSION AX268596

VERSION AX268596.1 GI:16541710

KEYWORDS

SOURCE Rattus sp.

```
ORGANISM
         Rattus sp.
         Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae;
REFERENCE
         1
         Cochran, S.W., Paterson, G.Y., Ohashi, Y.W., Morris, B.Y. and
 AUTHORS
         Pratt, J.Y.
         Schizophrenia related genes
 TITLE
         Patent: WO 0175440-A 15 11-OCT-2001;
 JOURNAL
         WELFIDE CORPORATION (JP)
                Location/Qualifiers
FEATURES
                1. .2697
   source
                /organism="Rattus sp."
                /mol type="unassigned DNA"
                /db xref="taxon:10118"
ORIGIN
 Query Match
                   81.8%; Score 2252.2; DB 6; Length 2697;
 Best Local Similarity
                   89.7%; Pred. No. 0;
 Matches 2419; Conservative
                         0; Mismatches 278;
                                          Indels
                                                  0:
                                                     Gaps
                                                           0:
        46 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 105
Qy
          1 ATGGCCGTCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60
Db
       106 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 165
Qу
          61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120
Db
       166 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 225
Qу
          121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180
Db
       226 GTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 285
Qу
          181 GTGTTGTTGGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240
Db
       286 TGGGTGCCCAGGTGGACCACGTGATCGAGCGCAGCAGACGGGAGCAGTGGGCTGCCC 345
Qy
          Db
       346 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 405
Qу
          301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360
Db
       406 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 465
Qу
          361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420
Db
       466 TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG 525
Qу
                        11111111 11 111
       421 TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG 480
Db
       526 TCCCTGGAGCAGGGCATCGTGCCGTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG 585
Qy
          Db
       481 TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCCAGAAGGAATCCCCCCAGCTGAG 540
```

| Qу | 586  | GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATC     | 645  |
|----|------|--|------|
| Db | 541  | GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCT                    | 600  |
| Qу | 646  | ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC | 705  |
| Db | 601  | ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC | 660  |
| Qу | 706  | TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC | 765  |
| Db | 661  | TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT | 720  |
| Qу | 766  | GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCC     | 825  |
| Db | 721  |  | 780  |
| Qу | 826  | GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC | 885  |
| Db | 781  | GGCTGGCAGAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC  | 840  |
| Qу | 886  | TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC | 945  |
| Db | 841  | TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGGAGC | 900  |
| Qу | 946  | TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT | 1005 |
| Db | 901  | TGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGCCGC | 960  |
| Qу | 1006 | GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCCAGGGCACTGACCTGGAC | 1065 |
| Db | 961  | GAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTGGAC | 1020 |
| Qу | 1066 | ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC | 1125 |
| Db | 1021 | ACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTGGCT | 1080 |
| Qу | 1126 | CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGCTGCTGCTGCTGCTCATC       | 1185 |
| Db | 1081 | CTCTACATCGGCCTTGTCGCTGTGGCTGTGCTCTTCTTGCTGTTGCTGGCCCTTGGA    | 1140 |
| QΨ | 1186 | CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC | 1245 |
| Db | 1141 | CTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATCCTC | 1200 |
| Qy | 1246 | ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC | 1305 |
| Db | 1201 | ACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTGCTC | 1260 |
| Qу | 1306 | ACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG    | 1365 |
| Db | 1261 | ACCATCCAGCCAGACCTCAGCACCACCACCTACCAGGGCAGTCTATGTTCGAGG       | 1320 |
| Qу | 1366 | CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT | 1425 |
| Db | 1321 |  | 1380 |
| Ωv | 1426 | CCCCCCCCCACACACACCACCACACCTCTCCCACCTCTCTCACCAC               | 1485 |

| Db   | 1381 | AGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTCTCC  | 1440 |
|------|------|---|------|
| ÒА   | 1486 | CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT  | 1545 |
| Db   | 1441 | CGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCCTAC  | 1500 |
| Qy   | 1546 | GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC  | 1605 |
| Db , | 1501 | GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTCCTC  | 1560 |
| Qy . | 1606 | ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG   | 1665 |
| Db   | 1561 | ATACCCCGGATGCCATCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCACAAG    | 1620 |
| Qу   | 1666 | CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC  | 1725 |
| Db   | 1621 | CCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTTAGC  | 1680 |
| Qу   | 1726 | TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG   | 1785 |
| Db   | 1681 | TGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGTGGA  | 1740 |
| Qу   | 1786 | GAGCCCAGCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG   | 1845 |
| Db   | 1741 | GAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGTTGG  | 1800 |
| Qу   | 1846 | GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG  | 1905 |
| Db   | 1801 | GAGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAGCTGGAG  | 1860 |
| Qу   | 1906 | GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC  | 1965 |
| Db   | 1861 | GCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGAGAGGCC  | 1920 |
| Qу   | 1966 | CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC  | 2025 |
| Db   | 1921 | CTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCCTGTACG  | 1980 |
| Qу   | 2026 | TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG  | 2085 |
| Db   | 1981 | TCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTCAAGGAG  | 2040 |
| Qу   | 2086 | GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC | 2145 |
| Db   | 2041 | GTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTCCTGCAC  | 2100 |
| Qy   | 2146 | TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG  | 2205 |
| Db   | 2101 | TTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCCCTGTGG  | 2160 |
| Qy   | 2206 | AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG  | 2265 |
| Db   | 2161 | AAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAACGGCACC  | 2220 |
| Qу   | 2266 | CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG  | 2325 |

| Db .  | 222  | 21 CAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGCGACCTG 2280   |  |
|---|------|--|--|
| Qy  | 232  | 26 GCCTGCAAGCTGTGGGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2385   |  |
| Db  | 228  | 31 GCCTGCAAGGTGTGGCTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATCAACTTC 2340   |  |
| Qу  | 238  | AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2445  |  |
| Db  | 234  | 11 AACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGGGGGGTC 2400   |  |
| Qу  | 244  | 6 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2505  |  |
| Db  | 240  | 1 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAGATCATC 2460  |  |
| Qу  | 250  | 06 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2565   |  |
| Db  | 246  | 51 GCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAGAAACTT 2520   |  |
| Qу  | 256  | 66 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTC 2625  |  |
| Db  | 252  | 21 CACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTC 2580   |  |
| Qу  | 262  | 26 AACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC   |  |
| Db  | 258  | 1 AACCTATGGGAGGCACGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCAGCAGCTGTG 2640  |  |
| QУ  | 268  | 6 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2742   |  |
| Db  | 264  | 1 GCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGTTGA 2697   |  |
| RESULT 9 RNU87305 LOCUS DEFINITI ACCESSION VERSION KEYWORDS | on · | RNU87305 2697 bp mRNA linear ROD 15-MAY-1997 Rattus norvegicus transmembrane receptor Unc5H1 mRNA, complete cds. U87305 U87305.1 GI:2055391        |  |
| SOURCE<br>ORGANI  | SM   | Rattus norvegicus (Norway rat) Rattus norvegicus   |  |
|   |      | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;                    |  |
| REFERENCE<br>AUTHORS  |      | Rattus.  1 (bases 1 to 2697)  Leonardo, E.D., Hinck, L., Masu, M., Keino-Masu, K., Ackerman, S.L. ar Tessier-Lavigne, M.                           |  |
| TITLE   |      | Vertebrate homologues of C. elegans UNC-5 are candidate netrin receptors   |  |
| JOURNAL<br>MEDLINE<br>PUBMED<br>REFERENCE<br>AUTHORS        |      | Nature 386 (6627), 833-838 (1997) 97271897 9126742 2 (bases 1 to 2697) Leonardo, E.D., Hinck, L., Masu, M., Keino-Masu, K. and Tessier-Lavigne, M. |  |
| TITLE<br>JOURNA   | L    | Direct Submission Submitted (28-JAN-1997) Anatomy, UCSF, 513 Parnassus, San Francisco, CA 94143-0452, USA  |  |

FEATURES

source

1. .2697

/organism="Rattus norvegicus"

/mol\_type="mRNA"

/db\_xref="taxon:10116"

/tissue\_type="brain and ventral spinal cord"

/dev\_stage="18 day embyo and 13 day embryo"

CDS

1. .2697

/codon\_start=1

/product="transmembrane receptor UNC5H1"

/protein\_id="AAB57678.1"

/db xref="GI:2055392"

translation="MAVRPGLWPVLLGIVLAAWLRGSGAQQSATVANPVPGANPDLLP/ HFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDSSSGLPTM EVRINVSRQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEV SLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTAN YTCVAKNIVARRRSTSAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPLN GGAFCEGONVOKTACATLCPVDGSWSSWSKWSACGLDCTHWRSRECSDPAPRNGGEEC RGADLDTRNCTSDLCLHTASCPEDVALYIGLVAVAVCLFLLLLALGLIYCRKKEGLDS DVADSSILTSGFOPVSIKPSKADNPHLLTIOPDLSTTTTTYQGSLCSRQDGPSPKFQL SNGHLLSPLGSGRHTLHHSSPTSEAEDFVSRLSTQNYFRSLPRGTSNMAYGTFNFLGG RLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPVVSCGPPGV LLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEESPSHLYYCQLEAGAC YVFTEOLGRFALVGEALSVAATKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKEVV QLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGT QQYLHCTFTLERINASTSDLACKVWVWQVEGDGQSFNINFNITKDTRFAELLALESEG GVPALVGPSAFKIPFLIROKIIASLDPPCSRGADWRTLAQKLHLDSHLSFFASKPSPT AMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC"

#### ORIGIN

Score 2252.2; DB 10; Length 2697; Query Match 81.8%; 89.7%; Best Local Similarity Pred. No. 0; Matches 2419; Conservative Mismatches 0; 0; 278; Indels 0; Gaps 46 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 105 Qy 1 ATGGCCGTCCGGCCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60 Db 106 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 165 Qу 61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGGCCAATCCAGTGCCCGGTGCCAACCCC 120 Db 166 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 225 Qу 121 GACCTGCTGCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180 Db 226 GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 285 Qy 181 GTGTTGTTGGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240 Db 286 TGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 345 Qу 241 TGGGTCCGCCAGGTCGATCACGTAATTGAACGCAGCACCCGACAGCAGCAGCGGATTGCCA 300 Db 346 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 405 Qy 301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360 Db

| Qу | 406  | GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC | 465  |
|----|------|--|------|
| Db | 361  | GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC | 420  |
| Qy | 466  | TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG | 525  |
| Db | 421  | TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG | 480  |
| Qу | 526  | TCCCTGGAGCAGGCATCGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG     | 585  |
| Db | 481  | TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCAGAAGGAATCCCCCCAGCTGAG  | 540  |
| Qу | 586  | GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC | 645  |
| Db | 541  | GTGGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCGATCCCAATGTGTACATC    | 600  |
| Qу | 646  | ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC | 705  |
| Db | 601  | ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC | 660  |
| Qу | 706  | TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC | 765  |
| Db | 661  | TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT | 720  |
| Qу | 766  | GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCC     | 825  |
| Db | 721  | GTGAACGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT     | 780  |
| Qу | 826  | GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC | 885  |
| Db | 781  | GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC | 840  |
| Qу | 886  | TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC | 945  |
| Db | 841  | TGTGAGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGGAGC  | 900  |
| Qy | 946  | TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT | 1005 |
| Db | 901  | TGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGCCGC | 960  |
| Qy | 1006 | GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGAC  | 1065 |
| Db | 961  | GAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTGGAC | 1020 |
| Qу | 1066 | ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC | 1125 |
| Db | 1021 | ACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTGGCT | 1080 |
| Qу | 1126 | CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATC | 1185 |
| Db | 1081 | CTCTACATCGGCCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTTTGGA       | 1140 |
| Qу | 1186 | CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC | 1245 |
| Db | 1141 | CTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATCCTC | 1200 |

|   | Qу   | 1246 | ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC  | 1305 |
|---|------|------|---|------|
|   | Db   | 1201 | ACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCACCTGCTC  | 1260 |
|   | Qy . | 1306 | ACCATCCAGCCGGACCTCAGCACCACCACCACCAGGGCAGTCTCTGTCCCCGG         | 1365 |
|   | Db   | 1261 |   | 1320 |
|   | Qу   | 1366 | CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT  | 1425 |
|   | Db   | 1321 |   | 1380 |
|   | Qy   | 1426 | GGCGGCCGCCACACTGCACCACACTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC     | 1485 |
|   | Db   | 1381 |   | 1440 |
|   | Qy   | 1486 | CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT  | 1545 |
|   | Db   | 1441 | CGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCCTAC  | 1500 |
|   | Qy   | 1546 | GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC  | 1605 |
|   | Db   | 1501 | GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTCCTC  | 1560 |
|   | Qy   | 1606 | ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG   | 1665 |
|   | Db   | 1561 |   | 1620 |
| ٠ | Qy   | 1666 | CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC  | 1725 |
|   | Db   | 1621 |   | 1680 |
|   | Qy   | 1726 | TGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG      | 1785 |
|   | Db   | 1681 | TGTGGGCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGTGGA   | 1740 |
|   | Qу   | 1786 | GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG  | 1845 |
|   | Db   | 1741 | GAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGTTGG  | 1800 |
|   | Qy   | 1846 | GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG  | 1905 |
|   | Db   | 1801 | GAGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAGCTGGAG  | 1860 |
|   | Qy   | 1906 | GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC  | 1965 |
|   | Db   | 1861 | GCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGAGAGGCC  | 1920 |
|   | Qy   | 1966 | CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC   | 2025 |
|   | Db   | 1921 |   | 1980 |
|   | Qу   | 2026 | TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG  | 2085 |
|   | Db   | 1981 | TCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTCAAGGAG  | 2040 |
|   | Ov   | 2086 | GTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC | 2145 |

| Db       | 2041  |   |
|----------|-------|---|
| Qy       | 2146  | TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2205 |
| Db       | 2101  | TTCAAAGACAGTTACCACAACCTACGTCTCTCCATCCACGACGTGCCCAGCTCCCTGTGG 2160 |
| Qy       | 2206  | AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2265 |
| Db       | 2161  | AAGAGCAAGCTACTTGTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAACGGCACC 2220 |
| Qу       | 2266  | CAGCGGTACTTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2325      |
| Db       | 2221  | CAGCAGTATCTGCACTGCACCTTCACCCTGGAGCGCATCAACGCCAGCACCAGCGACCTG 2280 |
| Qy .     | 2326  | GCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2385 |
| Db       | 2281  | GCCTGCAAGGTGTGGCAGGTGGAGGGAGATGGGCAGAGCTTCAACATCAACTTC 2340       |
| Qy       | 2386  | AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2445 |
| Db       | 2341  | AACATCACTAAGGACACAAGGTTTGCTGAATTGTTGGCTCTGGAGAGTGAAGGGGGGGTC 2400 |
| Qу       | 2446  | CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2505 |
| Db       | 2401  | CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAGATCATC 2460 |
| Qу       | 2506  | TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2565 |
| Db       | .2461 | GCCAGTCTGGACCCACCCTGCAGCCGGGGCGCCGACTGGAGAACTCTAGCCCAGAAACTT 2520 |
| Qу       | 2566  | CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2625 |
| Db       | 2521  | CACCTGGACAGCCATCTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTC 2580 |
| Qy       | 2626  | AACCTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC                     |
| Db       | 2581  | AACCTATGGGAGGCACGGCACCTCCCCAACGGCAACCTCGGCCAGCTGGCAGCAGCTGTG 2640 |
| Qу       | 2686  | GCTGGACTGGGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2742        |
| Db       | 2641  | GCCGGACTGGGCCAACCAGATGCTGGCCTCTTCACGGTGTCGGAGGCCGAGTGTTGA 2697    |
| RESULT 1 | 0     |   |

LOCUS BC058084 3844 bp mRNA linear ROD 21-OCT-2003

Mus musculus unc-5 homolog A (C. elegans), mRNA (cDNA clone DEFINITION

MGC:66671 IMAGE:6813463), complete cds.

ACCESSION BC058084

BC058084.1 GI:34784158 VERSION

KEYWORDS ' MGC.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

> Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE (bases 1 to 3844)

```
AUTHORS
            Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
            Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
            Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
            Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
            Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
            Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
            Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
            Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
            Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
            McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
            Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
            Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
            Fahey, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S.,
            Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
            Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
            Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
            Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
            Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
            Generation and initial analysis of more than 15,000 full-length
 TITLE
            human and mouse cDNA sequences
            Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
  JOURNAL
            22388257
 MEDLINE
            12477932
   PUBMED
            2 (bases 1 to 3844)
REFERENCE
            Strausberg, R.
 AUTHORS
            Direct Submission
 TITLE
  JOURNAL
            Submitted (08-SEP-2003) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            NIH-MGC Project URL: http://mgc.nci.nih.gov
  REMARK
COMMENT
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library Preparation: M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Genome Sequence Centre,
            BC Cancer Agency, Vancouver, BC, Canada
            info@bcqsc.bc.ca
            Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
            Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
            Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
            Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
            Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
            Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
            Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
            George Yang, Scott Zuyderduyn, Marco Marra.
            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 126 Row: b Column: 11
            This clone was selected for full length sequencing because it
```

passed the following selection criteria: matched mRNA gi: 23346570.

Location/Qualifiers

/mol type="mRNA"

/organism="Mus musculus"

1. .3844

**FEATURES** 

source

```
/strain="C57BL/6"
                /db xref="taxon:10090"
                /clone="MGC:66671 IMAGE:6813463"
                /tissue type="Brain, mouse, 13.5, 14.5, 16.5, 17.5 dpc"
                /clone lib="NIH BMAP FW0"
                /lab host="DH10B"
                /note="Vector: pYX-ASC"
                1. .3844
gene
                /gene="Unc5a"
                /db xref="LocusID:107448"
                /db_xref="MGI:894682"
CDS
                252. .2780
                /gene="Unc5a"
                /codon start=1
                /product="Unc5a protein"
                /protein id="AAH58084.1"
                /db xref="GI:34784159"
                /db xref="LocusID:107448"
                /db xref="MGI:894682"
                /translation="MAVRPGLWPALLGIVLTAWLRGSGAQQSATVANPVPGANPDLLP
                HFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTM
                EVRINVSRQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIAYLRKNFEQEPLAKEV
                SLEQGIVLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTPN
                YTCVAKNIVARRRSASAAVIVYVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGE
                ECRGADLDTRNCTSDLCLHTSSGPEDVALYIGLVAVAVCLILLLLVLVLIYCRKKEGL
                DSDVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKF
                QLSNGHLLSPLGSGRHTLHHSSPTSEAEDFVSRLSTQNYFRSLPRGTSNMAYGTFNFL
                GGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPP
                GVLLTRPVILAMDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEESPSHLYYCQLEAG
                ACYVFTEQLGRFALVGEALSVAATKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKE
                VVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWN
                GTQQYLHCTFTLERVNASTSDLACKVWVWQVEGDGQSFNINFNITKDTRFAEMLALES
                EGGVPALVGPSAFKIPFLIRQKIITSLDPPCSRGADWRTLAQKLHLDSHLSFFASKPS
                PTAMILNLWEARHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEC"
                714. .926
misc feature
                /gene="Unc5a"
                /note="IG; Region: Immunoglobulin"
                /db xref="CDD:smart00409"
                984. .1130
misc feature
                /gene="Unc5a"
                /note="TSP1; Region: Thrombospondin type 1 repeats"
                /db xref="CDD:smart00209"
                1566. .1877
misc feature
                /gene="Unc5a"
                /note="ZU5; Region: Domain present in ZO-1 and Unc5-like
                netrin receptors"
                /db xref="CDD:smart00218"
                249\overline{6}. .2684
misc feature
                /gene="Unc5a"
                /note="DEATH; Region: DEATH domain, found in proteins
                involved in cell death (apoptosis). Alpha-helical domain
                present in a variety of proteins with apoptotic functions.
                Some (but not all) of these domains form homotypic and
                heterotypic dimers"
                /db xref="CDD:smart00005"
```

Query Match 72.1%; Score 1983.6; DB 10; Length 3844; Best Local Similarity 84.3%; Pred. No. 0; Matches 2320; Conservative 0; Mismatches 264; Indels 168; Gaps 1;

| Qу   | 1   | CCGCGGGGCCCGCCCGCCCGCCCGCCCGCCCGCCCATGGCCGTCCGGCCC           | 60  |
|------|-----|--|-----|
| Db   | 207 | CCGCCGCGGGGCCCTAGTCCAGCCCGCTCGTCCGCCCGCGCCATGGCCGTCCGGCCC    | 266 |
| Qу   | 61  | GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC | 120 |
| Db   | 267 | GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCACTGCCTGGCTTCGTGGTTCGGGTGCC | 326 |
| Qy   | 121 | CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC | 180 |
| Db   | 327 | CAGCAGAGTGCCACAGTGCCAACCCAGTGCCTGCTGCCCAC                    | 386 |
| Qу   | 181 | TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC | 240 |
| Db   | 387 | TTCCTGGTAGAGCCGGAGGACGTGTACATTGTCAAGAACAAGCCCGTGCTGCTGGTGTGC | 446 |
| Qу   |     | AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG | 300 |
| Db · |     | AAGGCTGTGCCCGCCACCCAGATCTTCTTCAAGTGCAACGGGGAATGGGTTCGCCAGGTC | 506 |
| Qу   | 301 | GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC | 360 |
| Db   | 507 | GATCACGTCATTGAACGCAGCACTGACGGCAGCAGCGGATTGCCAACCATGGAGGTCCGG | 566 |
| Qу   | 361 | ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG | 420 |
| Db   | 567 | ATCAACGTATCAAGGCAGCAGGTCGAGAAAGTGTTTGGGCTGGAGGAGTACTGGTGCCAG | 626 |
| Qy   | 421 | TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC | 480 |
| Db   | 627 | TGTGTGGCATGGAGCTCCTCAGGAACCACCAAAAGCCAGAAGGCCTACATCCGGATTGCC | 686 |
| Qу   | 481 | AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC | 540 |
| Db   | 687 | TATTTGCGCAAGAACTTTGAGCAGGAGCCGCTGGCCAAGGAAGTGTCACTGGAGCAAGGC | 746 |
| Qу   | 541 | ATCGTGCTGCCCTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG      | 600 |
| Db   | 747 | ATTGTGCTACCTTGTCGCCCCCGGAAGGAATCCCCCCAGCTGAGGTGGAGTGGCTCCGA  | 806 |
| Qy   | 601 | AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC | 660 |
| Db   | 807 | AATGAGGACCTCGTGGACCCCTCGACCCCAATGTGTACATCACACGGGAGCACAGC     | 866 |
| QУ   | 661 | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC | 720 |
| Db   | 867 | CTAGTCGTGCGGCAGGCCCGCCTGGCCGACACTACACCTGCGTGGCCAAGAAC        | 926 |
| QУ   | 721 | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTG  | 780 |
| Db   | 927 | ATCGTGGCCCGTCGCCGAAGCGCCTCTGCGGCCGTCATTGTTTATGTG             | 974 |
| Qy   | 781 | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG | 840 |

| Db  | 975  |  | 974  |
|-----|------|--|------|
| Qy  | 841  | ${\tt AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT}$ | 900  |
| Db  | 975  |  | 974  |
| Qy  | 901  | GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGGAGC         |      |
| Db· | 975  | .  | 998  |
| Qу  | 961  | AAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCA         | 1020 |
| Db  | 999  | AAGTGGTCAGCCTGCGGGCTTGACTGCACCCACTGGCGGAGCCGGGAGTGCTCCGACCCA         | 1058 |
| Qу  | 1021 | GCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACC           | 1080 |
| Db  | 1059 | GCGCCCGCAACGGAGGTGAGGAGTGCCGGGGTGCTGACCTGGACACCCGCAACTGTACC          | 1118 |
| Qy  | 1081 | AGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTC         | 1140 |
| Db  | 1119 | AGTGACCTCTGCCTGCACACCTCTTCCGGCCCCGAGGACGTGGCTCTCTACATCGGCCTC         | 1178 |
| Qу  | 1141 | ATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGG         | 1200 |
| Db  | 1179 | GTCGCCGTGGCCGTGTGCCTCATCTTGCTGCTGCTGGTCCTCGTCCTCATCTACTGCCGC         | 1238 |
| Qy  | 1201 | AAGAAGGAGGGCTGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAG           | 1260 |
| Db  | 1239 | AAGAAGGAAGGACTCAGACGTGGCTGACTCATCCTTACCTCAGGCTTCCAG                  | 1298 |
| Qу  | 1261 | CCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGAC         | 1320 |
| Db  | 1299 | CCTGTCAGCATCAAGCCCAGCAAAGCAGACAATCCCCATCTGCTCACCATCCAACCGGAC         | 1358 |
| Qу  | 1321 | CTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGC            | 1380 |
| Db  | 1359 | CTCAGCACCACCACCACCTACCAGGGCAGCCTGTGTCCCCGGCAGGATGGACCCAGC            | 1418 |
| Qу  | 1381 | CCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACA         | 1440 |
| Db  | 1419 | CCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTGGGCAGTGGCCGCCATACG         | 1478 |
| Qу  | 1441 | CTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG         | 1500 |
| Db  | 1479 | CTGCACCACAGCTCCCCCACCTCTGAGGCTGAGGACTTCGTCTCCCGCCTCTCCACCCAA         | 1538 |
| Qу  | 1501 | AACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTC         | 1560 |
| Db  | 1539 | AACTACTTTCGTTCTCTGCCCCGCGGTACCAGCAACATGGCCTATGGGACCTTCAACTTC         | 1598 |
| Qу  | 1561 | CTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGATGCC         | 1620 |
| Db  | 1599 | CTCGGGGGCCGGCTGATGATCCCTAACACAGGAATCAGCCTCCTCATACCCCCGGACGCC         | 1658 |
| Qy  | 1621 | ATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGG         | 1680 |

|   | Db   | 1659 | ATCCCCCGAGGAAAGATCTACGAGATCTACCTCACTCTGCACAAGCCAGAAGACGTGAGG | 1718 |
|---|------|------|--|------|
|   | Qу   | 1681 | TTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGC | 1740 |
|   | Db   | 1719 | TTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCTATCGTTAGCTGTGGGCCCCCAGGA | 1778 |
|   | Qy   | 1741 | GTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGAC | 1800 |
|   | Db   | 1779 | GTCCTGCTCACCCGGCCAGTCATCCTTGCCATGGACCACTGCGGGGAGCCCAGTCCCGAC | 1838 |
|   | Qy   | 1801 | AGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCTGCAC | 1860 |
|   | Db   | 1839 | AGCTGGAGCCTGCAAAAAAGCAGTCCTGTGAGGGCAGCTGGGAGGACGTGCTGCAC     | 1898 |
|   | Qy   | 1861 | CTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTAC | 1920 |
|   | Db   | 1899 | CTTGGTGAGGAGTCGCCCTCTCATCTCTACTACTGCCAGCTGGAGGCCGGGGCCTGCTAT | 1958 |
|   | Qу   | 1921 | GTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCC | 1980 |
|   | Db   | 1959 | GTCTTTACCGAGCAGCTAGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCC | 2018 |
|   | Qy / | 1981 | GCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC          | 2040 |
|   | Db   | 2019 | ACCAAGCGCCTCAGGCTCCTTCTGTTTGCCCCTGTGGCCTGTACGTCCCTCGAGTACAAC | 2078 |
|   | Qу   | 2041 | ATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAG | 2100 |
|   | Db   | 2079 | ATCCGAGTGTACTGCCTGCACGACACCCACGATGCTCTCAAGGAGGTGGTGCAGCTGGAG | 2138 |
|   | Qy   | 2101 | AAGCAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTAC  | 2160 |
|   | Db   | 2139 | AAGCAGCTGGGTGGACAGCTGATCCAGGAGCCCCGTGTCCTGCACTTCAAAGACAGTTAC | 2198 |
|   | Qу   | 2161 | CACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTT | 2220 |
|   | Db   | 2199 | CACAACCTACGTCTGTCCATCCACGACGTGCCCAGCTCCCTGTGGAAGAGCAAGCTCCTT | 2258 |
|   | Qy . | 2221 | GTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCAC | 2280 |
| • | Db   | 2259 | GTCAGCTACCAGGAGATCCCTTTTTACCACATCTGGAATGGCACTCAGCAGTATCTGCAC | 2318 |
|   | Qy   | 2281 | TGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGG | 2340 |
|   | Db   | 2319 | TGCACCTTCACCCTGGAGCGCGTCAATGCCAGCACCAGCGACCTGGCCTGCAAGGTGTGG | 2378 |
|   | Qy   | 2341 | GTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGAC | 2400 |
|   | Db   | 2379 | GTGTGGCAGGTGGAGGAGATGGACAGAGCTTCAACATCAACTTTAACATCACTAAGGAC  | 2438 |
|   | Qy   | 2401 | ACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGC | 2460 |
|   | Db   | 2439 | ACGAGGTTTGCTGAAATGCTGGCTCTGGAGAGTGAAGGGGGGGTCCCAGCCCTGGTGGGC | 2498 |
|   | Qу   | 2461 | CCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCA | 2520 |
|   | Db   | 2499 | CCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAAAAGATCATTACCAGCCTGGACCCA | 2558 |

•

```
2521 CCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCAT 2580
Qv.
            2559 CCCTGCAGCCGGGGCGCCGACTGGCGAACTCTAGCCCAGAAACTTCACCTGGACAGCCAT 2618
Db.
        2581 CTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGGGAGGCG 2640
Qу
            2619 CTTAGCTTCTTTGCCTCCAAGCCCAGCCCTACAGCCATGATCCTCAACCTATGGGAGGCG 2678
Db
        2641 CGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAG 2700
Qу
            2679 CGGCACTTCCCCAACGGCAACCTCGGCCAGCTGGCCGCAGCTGTGGCCGGACTGGGCCAG 2738
Db
        Qу
            Db
RESULT 11
BC009333
           BC009333
                                2688 bp
                                                          PRI 25-MAR-2004
                                          mRNA
                                                  linear
LOCUS
          Homo sapiens unc-5 homolog A (C. elegans), mRNA (cDNA clone
DEFINITION
           IMAGE: 4126760), partial cds.
ACCESSION
           BC009333
VERSION
          BC009333.2 GI:40226527
KEYWORDS
SOURCE
          Homo sapiens (human)
          Homo sapiens
 ORGANISM
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
             (bases 1 to 2688)
REFERENCE
           Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 AUTHORS
           Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
          Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
           Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
           Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
           Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
           Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
           Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
          Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
          McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
          Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
          Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
           Fahey, J., Helton, E., Ketteman, M., Madan, A., Rodrigues, S.,
           Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
           Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
           Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
           Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
           Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
           Generation and initial analysis of more than 15,000 full-length
 TITLE
           human and mouse cDNA sequences
           Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 JOURNAL
           12477932
  PUBMED
             (bases 1 to 2688)
REFERENCE
 AUTHORS
           Strausberg, R.
 TITLE
           Direct Submission
 JOURNAL
           Submitted (12-JUN-2001) National Institutes of Health, Mammalian
```

```
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK
            NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT
            On Dec 19, 2003 this sequence version replaced gi:14424611.
            Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC),
            Gaithersburg, Maryland;
            Web site: http://www.nisc.nih.gov/
            Contact: nisc mgc@nhgri.nih.gov
            Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
            Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
            Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
            Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
            Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D.,McCloskey,J.C.,
            McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
            Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
            Young, A., Zhang, L.-H. and Green, E.D.
            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAL Plate: 26 Row: g Column: 22.
FEATURES
                     Location/Qualifiers
                     1. .2688
     source
                     /organism="Homo sapiens"
                     /mol type="mRNA"
                     /db xref="taxon:9606"
                     /clone="IMAGE:4126760"
                     /tissue type="Brain, neuroblastoma"
                     /clone lib="NIH MGC 19"
                     /lab host="DH10B-R"
                     /note="Vector: pOTB7"
                     <1. .2688
     gene
                     /gene="UNC5A"
                     /note="synonym: UNC5H1"
                     /db xref="LocusID:90249"
                     /db xref="MIM:607869"
    CDS
                     <1. .1627
                     /gene="UNC5A"
                     /codon start=2
                     /product="UNC5A protein"
                     /protein id="AAH09333.2"
                     /db xref="GI:40226528"
                     /db xref="LocusID:90249"
                     /db xref="MIM:607869"
                     translation="DVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILT/
                     SGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPL
                     GGGRHTLHHSSPTSEAEEFVSRLSTQNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGI
                     SLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILA
                     MDHCGEPSPDSWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGR
                     FALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGO
                     LIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFT
```

Gene Collection (MGC), Cancer Genomics Office, National Cancer

LERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPS AFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEA RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC"

misc feature

413. .724

/gene="UNC5A"

/note="ZU5; Region: Domain present in ZO-1 and Unc5-like

netrin receptors"

/db xref="CDD:smart00218"

misc\_feature

1343. .1594

/gene="UNC5A"

/note="DEATH; Region: DEATH domain, found in proteins involved in cell death (apoptosis). Alpha-helical domain present in a variety of proteins with apoptotic functions. Some (but not all) of these domains form homotypic and

heterotypic dimers"

/db xref="CDD:smart00005"

## ORIGIN

|         | ocal S | 59.4%; Score 1635.4; DB 9; Length 2688; Similarity 99.9%; Pred. No. 2.4e-264; 6; Conservative 0; Mismatches 1; Indels 0; Gaps | 0;         |
|---------|--------|---|------------|
| Macche. | 5 1050 | o, conservative o, mismatches i, indeis o, daps   | 0,         |
| Qy      |        | GGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCTGCTGCTGCTGCTGCTGCTGCTG  | 1175<br>60 |
| Db      | 1      | GGACGIGGCCCICIAIGIGGGCCICAICGCCGIGGCCGICIGCIGGICCIGCIGCI  | 00         |
| Qy      | 1176   | TGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTC  | 1235       |
| Db      | 61     | TGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGGCTGGACTCAGATGTGGCTGACTC   | 120        |
| QУ      | 1236   | GTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCC  | 1295       |
| Db      | 121    | GTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCC  | 180        |
| Qу      |        |   | 1355       |
| Db      |        | CCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCACCACCAGGGCAGTCT  |            |
| Qу      | •      | CTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAG  |            |
| Db      |        | CTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAG  |            |
| Qу      |        | CCCCTGGGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGA  |            |
| Db      |        | CCCCTGGGTGGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGA  |            |
| Qу      | 1476   | GTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAA  | 1535       |
| Db      | 361    | GTTCGTCTCCCGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAA  | 420        |
| Qy      | 1536   | CATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTAT  | 1595       |
| Db      | 421    | CATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAAT  | 480        |
| Qу      | 1596   | CAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCAC   | 1655       |
| Db      | 481    | CAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCAC   | 540        |

|   | Qy | 1656 | GCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTG                    | 1715 |
|---|----|------|---|------|
|   | Db | 541  | GCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTG                    | 600  |
|   | Qy | 1716 | CATCGTTAGCTGTGGACCCCCTGGCGTCCTCCTCACCCGGCCAGTCATCCTGGCTATGGA  | 1775 |
|   | Db | 601  | CATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGA  | 660  |
|   | Qy | 1776 | CCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGA  | 1835 |
| , | Db | 661  | CCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGA  | 720  |
|   | Qу | 1836 | GGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTG  | 1895 |
|   | Db | 721  | GGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTG  | 780  |
|   | Qу | 1896 | CCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGT  | 1955 |
|   | Db | 781  | CCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGT  | 840  |
|   | Qу | 1956 | GGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGT  | 2015 |
|   | Db | 841  | GGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGT  | 900  |
|   | Qy | 2016 | GGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC                  | 2075 |
|   | Db | 901  | GGCCTGCACCTCCGAGTACAACATCCGGGTCTACTGCCTGC                     | 960  |
|   | Qy | 2076 | ACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACG | 2135 |
|   | Db | 961  | ACTCAAGGAGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCACG    | 1020 |
|   | Qy | 2136 | GGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC               | 2195 |
|   | Db | 1021 | GGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC               | 1080 |
|   | Qy | 2196 | CTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTG  | 2255 |
|   | Db | 1081 | CTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTG  | 1140 |
|   | Qy | 2256 | GAATGGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCAC   | 2315 |
|   | Db | 1141 | GAATGGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCAC   | 1200 |
|   | Qy | 2316 | TAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAG  | 2375 |
|   | Db | 1201 | TAGTGACCTGCCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGCCAGAGCTTCAG      | 1260 |
|   | Qу | 2376 | CATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGA  | 2435 |
|   | Db | 1261 | CATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGA  | 1320 |
|   | Qу | 2436 | AGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCA  | 2495 |
|   | Db | 1321 | AGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCA  | 1380 |

```
2496 GAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGC 2555
Qy
           Db
       1381 GAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGC 1440
       2556 CCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGC 2615
Qy
           1441 CCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGC 1500
Db
       Qу
           Db
       2676 TGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGA 2735
Qy
           1561 TGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGA 1620
Db
       2736 GTGCTGAGGCCGGCCAG 2752
Qу
           1111111111111
Db
       1621 GTGCTGAGGCCGGCCAG 1637
RESULT 12
BD057525
LOCUS
         BD057525
                             1787 bp
                                      DNA
                                             linear
                                                    PAT 27-AUG-2002
DEFINITION
         Netrin receptors.
ACCESSION
         BD057525
VERSION
          BD057525.1 GI:22603131
KEYWORDS
          JP 2001505062-A/2.
SOURCE
          synthetic construct
 ORGANISM
         synthetic construct
          other sequences; artificial sequences.
REFERENCE
            (bases 1 to 1787)
 AUTHORS
         Lavigne, M.T., Leonardo, D.E., Hinck, L., Masu, M. and Masu, K.K.
 TITLE
         Netrin receptors
 JOURNAL
         Patent: JP 2001505062-A 2 17-APR-2001;
         THE REGENTS OF THE UNIV OF CALIFORNIA
COMMENT
          ΡN
              JP 2001505062-A/2
          PD
              17-APR-2001
              19-FEB-1998 JP 1998536840
          PF
              19-FEB-1997 US
         PR
                            08/808982
          PΙ
             MARC TESSIER LAVIGNE, DAVID E LEONARDO, LINDSAY HINCK, MASAYUKI
          PΙ
              MASU,
         PΙ
              KAZUKO KEINO MASU
              C07K1/00,C07K14/00,C07K17/00,C07H21/02,C07H21/04,G01N33/53 CC
          Strandedness: Double;
          CC
              Topology: Linear;
         FH
                           Location/Qualifiers.
              Key
                 Location/Qualifiers
FEATURES
                 1. .1787
    source
                 /organism="synthetic construct"
                 /mol type="genomic DNA"
                 /db xref="taxon:32630"
ORIGIN
 Query Match
                     56.8%; Score 1562.4; DB 6;
                                              Length 1787;
 Best Local Similarity
                     98.5%; Pred. No. 4.3e-252;
 Matches 1661; Conservative 0; Mismatches
                                         16; Indels
                                                     9; Gaps
                                                               8;
```

| Q:<br>Dl         | •                                     | GCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCT                  |      |
|------------------|---------------------------------------|---|------|
| Q                | 7 1130                                | . ATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGCTGCTGCTGCTTGTCCTCATCCTCG                   | 1189 |
| D)               | ·<br>60                               |   | 119  |
| Q:               | 1190                                  | TTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCT                  | 1249 |
| Dl               | 120                                   |   | 179  |
| Q                | 1250                                  | CAGGCTTCCAGCCCGTCAGCATC-AAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACC                  | 1308 |
| Dl               | 180                                   |   | 239  |
| Q                | 1309                                  | ATCCAGCCGGACCTCAGCACCACCACCACCACCAGGGCAGTCTCTGTCCCCGGCAG                      | 1368 |
| D}               | 240                                   | ATCCAGCCGGACCTCAGCACCACCACCACCAGGGCAGTCTCTGTCCCCGGCAG                         | 299  |
| Q                | 1369                                  | GATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC                  | 1428 |
| Dl               | 300                                   | GATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGC                  | 359  |
| ,Q <u>:</u>      | 1429                                  | GGCCGCCACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC                    | 1488 |
| Dl               | 360                                   | GGCCGCCACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGC                    | 419  |
| Q                | 1489                                  | CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGG                  | 1548 |
| Dł               | 420                                   | CTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGG                  | 479  |
| Q <u>r</u>       |                                       | ACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATC                  |      |
| Dł               |                                       | ACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATC                  |      |
| Q <u>y</u>       |                                       | CCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG                  |      |
| Di               | , , , , , , , , , , , , , , , , , , , | CCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCG                  |      |
| Q <u>y</u>       |                                       | GAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT                  |      |
| Di               |                                       | GAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGT  CCACGCCCCCCCCCC |      |
| Q <u>y</u><br>Dl |                                       | GGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAG                      |      |
|                  |                                       | CCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAG                  |      |
| Q <u>y</u><br>Dl |                                       | CCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGCAGCTGGGAG                   |      |
| Q <sub>3</sub>   |                                       | GATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCC                  |      |
| Dl               |                                       |   |      |
|                  |                                       |   |      |

| Qу |      | AGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTC | 1968 |
|----|------|--|------|
| Db |      |  | 897  |
| Qу |      | AGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCC | 2028 |
| Db |      |  | 957  |
| Qy |      | CTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTG | 2088 |
| Db |      |  | 1017 |
| Qу | 2089 | GTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTC | 2148 |
| Db | 1018 |  | 1076 |
| Qy |      | AAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAG | 2208 |
| Db |      |  | 1134 |
| Qy |      | AGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAG | 2268 |
| Db |      |  | 1194 |
| Qy |      | CGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCC | 2328 |
| Db |      |  | 1254 |
| Qy |      | TGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAAC | 2388 |
| Db |      |  | 1314 |
| Qу | 2389 | ATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCA | 2448 |
| Db | 1315 |  | 1374 |
| Qу |      | GCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCC | 2508 |
| Db |      |  | 1434 |
| Qy |      | AGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCAC | 2568 |
| Db |      | AGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCAC | 1494 |
| Qy | 2569 | CTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAAC | 2628 |
| Db | 1495 | CTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAAC | 1554 |
| Qy | 2629 | CTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC                   | 2688 |
| Db | 1555 | CTGTGGGAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGC                   | 1614 |
| Qy | 2689 | GGACTGGGCCAGCCAGACGCTGGCCTC-TTCACAGTG-TCGGAGGCTGAGTGCTGAGGCC | 2746 |
| Db | 1615 |  | 1674 |
| Qу | 2747 | GGCCAG 2752  |      |
| ,  |      |  |      |
|    |      |  | •    |
|    |      | •  |      |
|    |      |  |      |

|||||| Db 1675 GGCCAG 1680

```
RESULT 13 ·
AX054976
                                                    PAT 13-JAN-2001
LOCUS
         AX054976
                             9700 bp
                                      DNA
                                             linear
         Sequence 91 from Patent WO0073328.
DEFINITION
ACCESSION
         AX054976
VERSION
         AX054976.1 GI:12228344
KEYWORDS
SOURCE
         synthetic construct
         synthetic construct
 ORGANISM
         other sequences; artificial sequences.
REFERENCE
 AUTHORS
         van Criekinge, W., Roelens, I., Bogaert, T. and Verwaerde, P.
         Unc-5 constructs and screening methods
 TITLE
         Patent: WO 0073328-A 91 07-DEC-2000;
 JOURNAL
         Devgen NV (BE)
FEATURES
                 Location/Oualifiers
                 1. .9700
    source
                 /organism="synthetic construct"
                 /mol type="unassigned DNA"
                 /db xref="taxon:32630"
                 /note="plasmid pGC1037"
ORIGIN
                           Score 1304.2; DB 6;
 Query Match
                     47.4%;
                                             Length 9700;
 Best Local Similarity
                           Pred. No. 6.8e-209;
                     99.8%;
 Matches 1306; Conservative
                           0; Mismatches
                                         3;
                                            Indels
                                                               0;
                                                     0; Gaps
       1437 CACACTGCACCACGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCAC 1496
Qу
           7 CACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCAC 66
Db
       1497 CCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAA 1556
Qу
           67 CCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAA 126
Db
       1557 CTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGA 1616
Qу
           127 CTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTCATCCCCCAGA 186
Db
       1617 TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGT 1676
Qy
           187 TGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGT 246
Db
       1677 GAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC 1736
Qу
           247 GAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCC 306
Db
       1737 TGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCC 1796
Qу
           307 TGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCC 366
Db
Qy
       1797 TGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCT 1856
```

| Db  | 367  | TGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCT       | 426  |
|-----|------|--|------|
| Qy  | 1857 | GCACCTGGGCGAGGAGGCCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTG        | 1916 |
| Db  | 427  | GCACCTGGGCGAGGGGCCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTG         | 486  |
| Qy  | 1917 | CTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGC       | 1976 |
| Db  | 487  | $\tt CTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGC$ | 546  |
| Qу  | 1977 | TGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC            | 2036 |
| Db  | 547  | TGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC            | 606  |
| Qy  | 2037 | CAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCT       | 2096 |
| Db  | 607  | CAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCT       | 666  |
| Qy  | 2097 | GGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAG       | 2156 |
| Db  | 667  | GGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAG       | 726  |
| Qy  | 2157 | TTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCT       | 2216 |
| Db  | 727  | TTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCT       | 786  |
| Qу  | 2217 | CCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTT       | 2276 |
| Db  | 787  | CCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTT       | 846  |
| Qy  | 2277 | GCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCT       | 2336 |
| Db  | 847  | GCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCT       | 906  |
| Qy  | 2337 | GTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACATCACCAA             | 2396 |
| Db  | 907  | GTGGGTGTGGCAGGGGGGGGGGGGGGGGGGGGGGGGGGG                            | 9.66 |
| Qу  | 2397 | GGACACAAGGTTTGCTGAGCTGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGT           | 2456 |
| Dp. | 967  | GGACACAAGGTTTGCTGAGCTGCTCTGGAGAGTGAAGCGGGGGTCCAAGCCCTGGT           | 1026 |
| Qу  | 2457 | GGGCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGA        | 2516 |
| Db  | 1027 | GGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGA       | 1086 |
| Qy  | 2517 | CCCACCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAG        | 2576 |
| Db  | 1087 | CCCACCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAG        | 1146 |
| Qy  | 2577 | CCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGGGA        | 2636 |
| Db  | 1147 | CCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGGGA        | 1206 |
| Qy  | 2637 | GGCGCGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTG                           | 2696 |
| Db  | 1207 |  | 1266 |

```
2697 CCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGC 2745
Qy
              Db
         1267 CCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGTC 1315
RESULT 14
MMU72634
LOCUS
            MMU72634
                                    9299 bp
                                              mRNA
                                                       linear
                                                                ROD 13-MAY-1997
DEFINITION
           Mus musculus rostral cerebellar malformation protein (rcm) mRNA,
            complete cds.
            U72634
ACCESSION
            U72634.1 GI:2088526
VERSION
KEYWORDS
            Mus musculus (house mouse)
SOURCE
  ORGANISM
           Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
               (bases 1 to 9299)
  AUTHORS
            Ackerman, S.L., Kozak, L.P., Przyborski, S.A., Rund, L.A., Boyer, B.B.
            and Knowles, B.B.
            The mouse rostral cerebellar malformation gene encodes an
  TITLE
            UNC-5-like protein
            Nature 386 (6627), 838-842 (1997)
  JOURNAL
            97271898
  MEDLINE
            9126743
   PUBMED
REFERENCE
               (bases 1 to 9299)
  AUTHORS
            Ackerman, S.L., Kozak, L.P., Rund, L.A. and Knowles, B.B.
  TITLE
            Direct Submission
            Submitted (25-SEP-1996) The Jackson Laboratory, 600 Main Street,
  JOURNAL
            Bar Harbor, ME 04609, USA
FEATURES
                     Location/Qualifiers
                     1. .9299
     source
                     /organism="Mus musculus"
                     /mol type="mRNA"
                     /strain="C57B6/SJL"
                     /db xref="taxon:10090"
                     /chromosome="3"
                     /map="between Amy1 and Mpmv9"
                     1. .9299
     gene
                     /gene="rcm"
                     151. .2946
     CDS
                     /gene="rcm"
                     /function="transmembrane receptor"
                     /note="similar to C. elegans UNC-5"
                     /codon start=1
                     /product="rostral cerebellar malformation protein"
                     /protein id="AAB54103.1"
                     /db xref="GI:2088527"
                     translation="MRKGLRATAARCGLGLGYLLQMLVLPALALLSASGTGSAAQDDE/
                     FFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQ
                     KDHVVDERVDETSGLIVREVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYV
                     RIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEDIIDPAEDRNFYI
                     TIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRC
                     GRGYQKRTRTCTNPAPLNGGAFCEGQSVQKIACTTLCPVDGRWTSWSKWSTCGTECTH
```

WRRRECTAPAPKNGGKDCDGLVLQSKNCTDGLCMQAAPDSDDVALYVGIVIAVTVCLA ITVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQDLLAVPPDLTSAAAMYR GPVYALHDVSDKIPMTNSPILDPLPNLKIKVYNSSGAVTPQDDLAEFSSKLSPQMTQS LLENEALNLKNQSLARQTDPSCTAFGTFNSLGGHLIIPNSGVSLLIPAGAIPQGRVYE MYVTVHRKENMRPPMEDSQTLLTPVVSCGPPGALLTRPVILTLHHÇADPSTEDWKIQL KNQAVQGQWEDVVVVGEENFTTPCYIQLDAEACHILTENLSTYALVGQSTTKAAAKRL KLAIFGPLCCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKALHFKGSIHN LRLSIHDIAHSLWKSKLLAKYQEIPFYHIWSGSQRNLHCTFTLERLSLNTVELVCKLC VRQVEGEGQIFQLNCTVSEEPTGIDLPLLDPASTITTVTGPSAFSIPLPIRQKLCSSL DAPQTRGHDWRMLAHKLNLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLE EMGRHETVVSLAAEGQY"

## ORIGIN

| Query M |     | 36.0%; Score 992; DB 10; Length 9299;<br>Similarity 62.7%; Pred. No. 1.5e-156; |     |
|---------|-----|--|-----|
|         |     | 0; Conservative 0; Mismatches 925; Indels 63; Gaps                             | 5;  |
| Qy      | 149 | TGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACA                   | 208 |
| Db      | 308 |  | 367 |
| Qу      | 209 | TCGTCAAGAACAAGCCAGTGCTGCTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCT                    | 268 |
| Db      | 368 |  | 427 |
| Qу      | 269 | TCAAGTGCAACGGGGAGTGGGTGCGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACG                   | 328 |
| Db      | 428 |  | 487 |
| Qу      | 329 | GGAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGA                   | 388 |
| Db      | 488 |  | 547 |
| Qу      | 389 | AGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCA                   | 448 |
| Db      | 548 |  | 607 |
| Qу      | 449 | CCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGC                   | 508 |
| Db      | 608 | CGAAGAGTCGGAAGGCATACGTGCGCATTGCGTATCTGCGGAAGACATTCGAGCAGGAAC                   | 667 |
| Qу      | 509 | CGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGG                   | 568 |
| Db      | 668 | CCTTGGGAAAGGAAGTGTCCTTGGAGCAGGAAGTCTTACTCCAGTGTCGGCCACCTGAAG                   | 727 |
| Qу      | 569 | GCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGG                   | 628 |
| Db      | 728 |  | 787 |
| Qу      | 629 | ACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTG                   | 688 |
| Db      | 788 |  | 847 |
| Qу      | 689 | ACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCG                   | 748 |
| Db      | 848 |  | 907 |
| Qу      | 749 | CTGCTGTCATCGTCTACGTGAACGGTGGGTCGACGTGGACCGAGTGGTCCGTCTGCA                      | 808 |

| Db  | 908  | CCACTGTCATCGTGTATGTTAATGGTGGCTGGTCCACCTGGACAGAGTGGTCTGTGTA           | 967  |
|-----|------|--|------|
| Qy  | 809  | GCGCCAGCTGTGGGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTC           | 868  |
| Db  | 968  | ACAGCCGCTGTGGGCGAGGATATCAGAAACGCACAAGAACCTGCACCAACCCAGCCCCAC         | 1027 |
| Qу  | 869  | TCAACGGGGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGT          | 928  |
| Db  | 1028 | TCAATGGTGGGGCCTTCTGTGAGGGGCAGAGTGTGCAGAAAATAGCATGCACTACGTTAT         | 1087 |
| QУ  | 929  | GCCCAGTAGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCA         | 988  |
| Db  | 1088 | GTCCAGTGGATGGTAGGTGGACTTCATGGAGCAAATGGTCAACCTGTGGGACTGAATGCA         | 1147 |
| Qy  | 989  | CCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCC         | 1048 |
| Db  | 1148 | CCCACTGGCGCAGGAGGGAGTGTACAGCACCAGCCCCCAAGAACGGGGGTAAGGACTGTG         | 1207 |
| Qy  | 1049 | AGGGCACTGACCTGGACACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTG         | 1108 |
| Db  | 1208 | ATGGCCTGGTCCTCCAATCCAAGAACTGCACTGATGGGCTGTGCATGCA                    | 1267 |
| Qу  | 1109 | GCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCC            | 1165 |
| Db  | 1268 | ACTCAGATGATGTGGCTCTCTACGTGGGGATTGTGATCGCTGTAACAGTCTGTCT              | 1327 |
| QУ  | 1166 | TGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATG         | 1225 |
| Db  | 1328 | TCACTGTTGTGGTGGCCCTGTTTGTGTATCGGAAGAACCACCGTGACTTTGAGTCTGACA         | 1387 |
| Qy  | 1226 | TGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAG         | 1285 |
| Db  | 1388 | TCATTGACTCCTCAGCACTCAATGGCGGCTTTCAGCCTGTGAACATCAAGG                  | 1438 |
| Øλ. | 1286 | CAGACAACCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACC          | 1345 |
| Db  | 1439 | CTGCCAGACAAGATCTCCTGGCTGTCCCCCCTGACCTCACCTCAGCTGCAGCCATGTACA         | 1498 |
| Qу  | 1346 | AGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAA              | 1400 |
| Db  | 1499 | GGGGACCTGTCTATGCTCTGCATGATGTCTCAGACAAAATCCCAATGACCAACTCTCCAA         | 1558 |
| Qу  | 1401 | TGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACACTGCACCACAGCTCTCCCAC         | 1460 |
| Db  |      | ${\tt TTCTGGACCCACTACCCAACTTGAAAATCAAAGTGTACAACAGCTCAGGTGCTGTCACTC}$ |      |
| Qy  | 1461 | CTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCC                               | 1498 |
| Db  |      | CTCAGGATGACCTTGCCGAGTTCTCATCCAAACTGTCACCCCAGATGACCCAGTCCTTGC         |      |
| Qy  | 1499 | AGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCA                                 | 1534 |
| Db  | 1679 | TAGAGAATGAGGCCCTTAACCTGAAGAACCAGAGCCTCGCAAGACAGAC                    | 1738 |
| QУ  | 1535 | ACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTA         | 1594 |
| Db  | 1739 | GCACAGCATTTGGTACCTTCAACTCTCTTGGGGGTCACCTCATCATTCCTAATTCAGGAG         | 1798 |

|   | Qу   | 1595 | TCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCA   | 1654 |
|---|------|------|---|------|
|   | Db   | 1799 | TAAGCTTGCTGATTCCCGCTGGGGCCATTCCTCAGGGGAGAGTCTATGAAATGTATGT    | 1858 |
|   | Qу   | 1655 | CGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTG                   | 1714 |
|   | Db   | 1859 | CTGTACACAGGAAAGAAAATATGAGGCCCCCCATGGAAGACTCTCAGACCCTACTTACCC  | 1918 |
|   | Qу   | 1715 | CCATCGTTAGCTGTGGACCCCTGGCGTCCTCACCCGGCCAGTCATCCTGGCTATGG      | 1774 |
|   | Db   | 1919 | CTGTGGTGAGCTGTGGGGCCTCCTGGAGCTCTGCTGACCCGCCCTGTCATCCTCACTCTGC | 1978 |
|   | Qу   | 1775 | ACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCG  | 1834 |
|   | Db   | 1979 | ATCACTGTGCAGACCCCAGCACCGAGGACTGGAAGATCCAGCTCAAAAACCAGGCAGTGC  | 2038 |
|   | Qу   | 1835 | AGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACT  | 1894 |
|   | Db   | 2039 | AGGGACAATGGGAGGATGTTGTGGTGGTTGGGGAGGAGAACTTCACAACCCCCTGTTACA  | 2098 |
|   | Qу   | 1895 | GCCAGCTGGAGGCCAGTGCCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGG     | 1954 |
|   | Db   | 2099 | TTCAGCTGGATGCAGAGGCTTGCCATATCCTCACAGAGAACCTCAGTACCTATGCCCTGG  | 2158 |
|   | Qу   | 1955 | TGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGG  | 2014 |
|   | Db   | 2159 | TTGGGCAGTCCACCAAAGCAGCTGCCAAGCGTCTTAAACTGGCCATCTTTGGGCCCC     | 2218 |
|   | Qу   | 2015 | TGGCCTGCACCTCCCTCGAGTACAACATCCGGGTCTACTGCCTGC                 | 2074 |
| • | Db   | 2219 | TCTGCTGCTCTTCCCTGGAGTACAGCATTAGAGTCTACTGCCTGGATGACACACAGGATG  | 2278 |
|   | Qy   | 2075 | CACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCAC | 2134 |
|   | Db . | 2279 | CCCTGAAGGAAGTTCTACAACTGGAGAGGCCAAATGGGAGGACAGCTCCTAGAAGAACCCA | 2338 |
|   | Qу   | 2135 | GGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC              | 2194 |
|   | Db   | 2339 | AGGCTCTTCATTTTAAAGGCAGCATCCACAACCTGCGCCTGTCTATTCATGACATCGCCC  | 2398 |
| - | Qу   | 2195 | GCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCT  | 2254 |
|   | Db   | 2399 | ATTCCCTCTGGAAGAGCAAATTGCTGGCTAAGTATCAGGAAATTCCATTTTACCACATCT  | 2458 |
|   | Qу   | 2255 | GGAATGGCACGCAGCGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCA   | 2314 |
|   | Db   | 2459 | GGAGTGGCTCTCAAAGAAACCTCCACTGCACCTTCACTCTGGAAAGACTCAGCCTAAACA  | 2518 |
|   | QУ   | 2315 | CTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCA  | 2374 |
|   | Db   | 2519 | CAGTGGAACTGGTTTGCAAACTCTGTGTGCGGCAGGTTGAAGGAGAAGGGCAGATCTTCC  | 2578 |
|   | Qу   | 2375 | GCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTG  | 2434 |
|   | Db   | 2579 | AGCTCAACTGTACTGTCAGAGGAACCTACTGGCATCGACTTACCTCTCCTGGACCCTG    | 2638 |

```
2435 AAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGC 2494
Qу
                              Db
       2495 AGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGG 2554
Qу
                                                 111111 111
                      - 1
                                         1 11
       2699 AGAAGCTATGCAGCAGCCTGGATGCCCCTCAAACAAGAGGCCATGACTGGAGGATGCTGG 2758
Db
       2555 CCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAG 2614
Qу
            2759 CCCATAAACTCAACCTGGACAGGTACTTGAATTACTTTGCCACCAAATCGAGCCCAACTG 2818
Db
       Qу
            2819 GCGTAATCCTGGATCTTTGGGAAGCACAGAACTTCCCAGATGGAAACCTGAGCATGCTGG 2878
Db
       2675 CTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTG 2734
Qу
                                 1 1 1
                                                  - 1
                                                       1 11 1
                  2879 CAGCCGTCCTGGAAGAAATGGGAAGACATGAGACAGTGGTGTCCTTGGCAGCAGAAGGAC 2938
Db
       2735 AGTGCTGA 2742
Qy
            111 111
Db
       2939 AGTATTGA 2946
RESULT 15
AY187310
LOCUS
          AY187310
                              2962 bp
                                       mRNA
                                              linear
                                                     VRT 06-JUN-2003
         Gallus gallus UNC5-like protein 3 mRNA, complete cds.
DEFINITION
ACCESSION
          AY187310
VERSION
          AY187310.1 GI:31442350
KEYWORDS
          Gallus gallus (chicken)
SOURCE
 ORGANISM
          Gallus gallus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
          Phasianinae; Gallus.
          1 (bases 1 to 2962)
REFERENCE
          Guan, W. and Condic, M.L.
 AUTHORS
          Characterization of Netrin-1, Neogenin and cUNC-5H3 expression
 TITLE
          during chick dorsal root ganglia development
          Gene Expr. Patterns 3, 369-373 (2003)
 JOURNAL
            (bases 1 to 2962)
REFERENCE
          Guan, W. and Condic, M.L.
 AUTHORS
          Direct Submission
 TITLE
          Submitted (26-NOV-2002) Neurobiology & Anatomy, University of Utah,
 JOURNAL
          20 North, 1900 East, Salt Lake City, UT 84132-3401, USA
                 Location/Qualifiers
FEATURES
                 1. .2962
    source
                 /organism="Gallus gallus"
                 /mol type="mRNA"
                  /db xref="taxon:9031"
    CDS
                 1. .2796
                 /codon start=1
                 /product="UNC5-like protein 3"
                 /protein_id="AA067275.1"
                  /db xref="GI:31442351"
```

/translation="MGKGLEGTAARCGLGMGYLLHSVVLPALAVLGASRPGSAAQDDD FFHELPETFPSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQ KDHVVDERVDETSGLIVCEVSIEISROOVEELFGPEDYWCQCVAWSSAGTTKSRKAYV RIAYLRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEEVIDPVEDRNFYI TIDHNLIIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSACNSRC GRGFOKRTRTCTNPAPLNGGAFCEGONVOKIACTTLCPVDGKWTSWSKWSTCGTECTH WRRRECTAPAPKNGGKDCEGLVLQSKNCTDGLCMQAAPDSDDVALYVGIVIAVIVCLA  ${\tt ISVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNIKAARQDLLAVPPDLTSAAAMYR}$ GPVYALHDVSDKIPMTNSPILDPLPNLKIKVYNTSGAVTPQDELSDFSSKLSPQITQS LLENETLNVKNQSLARQTDPSCTAFGTFNSLGGHLVIPNSGVSLLIPAGAVPQGRVYE MYVTVHRKEGMRPPVEDSQTLLTPVVSCGPPGALLTRPVVLTMHHCAEPNMDDWQIQL KHQAGQGPWEDVVVVGEENFTTPCYIQLDPEACHILTETLSTYALVGQSITKAAAKRL KLAIFGPLSCSSLEYSIRVYCLDDTQDALKEVLQLERQMGGQLLEEPKTLHFKGSTHN LRLSIHDIAHSLWKSKLPAKYQEIPFYHIWSGCQRNLHCTFTLERFSLNTLELVCKLC VRQVEGEGQIFQLNCSVSEEPTGIDYPIMDSAGSITTIVGPNAFSIPLPIRQKLCSSL DAPQTRGHDWRMLAHKLKLDRYLNYFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLE EMGRHETVVSLAAEGNY"

#### ORIGIN

| Query Match 36.0%; Score 991; DB 5; Length 2962; Best Local Similarity 62.6%; Pred. No. 2.6e-156; |     |  |     |  |  |  |
|---|-----|--|-----|--|--|--|
| Matches   | 166 | 2; Conservative 0; Mismatches 930; Indels 63; Gaps           | 5;  |  |  |  |
| Qу  | 149 | TGCCTGGTGCCAACCCGGACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACA | 208 |  |  |  |
| Db  | 158 | TTCCTTCTGATCCTCCAGAGCCATTGCCCCACTTTCTCATTGAACCCGAAGAAGCTTACA | 217 |  |  |  |
| Qу  | 209 | TCGTCAAGAACAAGCCAGTGCTGCTTGTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCT | 268 |  |  |  |
| Db  | 218 | TCGTGAAAAACAAGCCTGTGAATCTGTACTGCAAAGCGAGCCCTGCCACGCAGATCTATT | 277 |  |  |  |
| Qy  | 269 | TCAAGTGCAACGGGGAGTGGGTGGCCAGGTGGACCACGTGATCGAGCGCAGCACAGACG  | 328 |  |  |  |
| Db  | 278 | TTAAGTGCAACAGTGAATGGGTTCATCAGAAGGATCATGTGGTGGATGAGAGAGTAGATG | 337 |  |  |  |
| Qу  | 329 | GGAGCAGTGGGCTGCCCACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGA | 388 |  |  |  |
| Db .  | 338 | AAACCTCTGGTCTGATCGTCTGTGAGGTGAGCATCGAGATTTCCCGCCAGCAGGTGGAAG | 397 |  |  |  |
| Qу  | 389 | AGGTGTTCGGGCTGGAGGAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCA | 448 |  |  |  |
| Db  | 398 | AGCTCTTTGGACCCGAGGACTACTGGTGCCAGTGTGTCGCCTGGAGCTCAGCTGGCACCA | 457 |  |  |  |
| Qу  | 449 | CCAAGAGTCAGAAGGCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGC | 508 |  |  |  |
| Db .  | 458 | CCAAGAGCCGCAAGGCCTACGTCCGCATTGCATATCTCAGAAAGACTTTTGAGCAGGAGC | 517 |  |  |  |
| Qу  | 509 | CGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGCATCGTGCTGCCCTGCCGTCCACCGGAGG | 568 |  |  |  |
| Db  | 518 | CGCTGGGGAAAGAAGTGTCCCTGGAGCAAGAGGTCCTGCTCCAGTGCCGTCCTCCTGAAG | 577 |  |  |  |
| Qу  | 569 | GCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGG | 628 |  |  |  |
| Db  | 578 | GCATTCCAGTAGCTGAGGTAGAGTGGCTGAAGAATGAAGAGGTGATCGATC          | 637 |  |  |  |
| Qу  | 629 | ACCCCAATGTATACATCACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTG | 688 |  |  |  |

|   | Db | 638         | ACCGAAATTTTTACATCACCATTGATCACAACCTGATCATCAAGCAAG              | 697  |
|---|----|-------------|---|------|
|   | Qy | 689         | ACACGGCCAACTACACCTGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCG  | 748  |
|   | Db | 698         | ACACGGCTAACTACACCTGTGTTGCCAAAAACATTGTGGCCAAAAGGAAAAGCACGACAG  | 757  |
|   | Qу | 749         | CTGCTGTCATCGTCAACGGTGGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCA        | 808  |
|   | Db | 758         | CAACTGTGATTGTCTATGTGAATGGAGGCTGGTCTACCTGGACCGAGTGGTCAGCGTGCA  | 817  |
|   | Qу | 809         | GCGCCAGCTGTGGGCGGCTGGCAGAAACGGAGCCGGAGCTGCACCCAACCCGGCGCCTC   | 868  |
|   | Db | 818         | ACAGCCGCTGTGGGAGAGGCTTCCAGAAGCGCACAAGGACCTGCACTAACCCTGCCCCAC  | 877  |
|   | Qу | 869         | TCAACGGGGCGCTTTCTGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGT   | 928  |
|   | Db | 878         | TCAATGGGGGGCCTTCTGCGAGGGGCAAAATGTTCAGAAAATAGCTTGCACCACCCTGT   | 937  |
|   | Qу | 929         | GCCCAGTAGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCA  | 988  |
|   | Db | 938         | GTCCAGTGGATGGCAAATGGACGTCCTGGAGCAAGTGGTCCACTTGTGGCACAGAGTGTA  | 997  |
|   | Qу | 989         | CCCACTGGCGGAGCCGTGAGTGCTCTGACCCAGCACCCCGCAACGGAGGGGAGGAGTGCC  | 1048 |
|   | Db | 998         | CCCACTGGCGCCGGAGGAGTGCACAGCTCCGGCCCCGAAGAATGGAGGCAAGGACTGTG   | 1057 |
|   | Qу | 1049        | AGGGCACTGACCTGGACACCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTG   | 1108 |
|   | Db | 1058        | AGGGACTGGTGCAGTCTAAGAACTGCACTGATGGGCTCTGCATGCA                | 1117 |
|   | Qу | 1109        | GCCCTGAGGACGTGGCCCTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCC     | 1165 |
|   | Db | 1118        | ACTCGGATGATGTTGCTCTCTACGTGGGGATTGTCATTGCTGTGATTGTGTGCCTGGCTA  | 1177 |
|   | Qy | 1166        | TGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGGAAGAAGGAGGGGGCTGGACTCAGATG | 1225 |
|   | Db | <b>1178</b> | TTTCTGTGGTTGTGGCCCTGTTTGTCTATCGCAAGAACCACCGTGACTTTGAGTCAGATA  | 1237 |
|   | Qу | 1226        | TGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAG  | 1285 |
| • | Db | 1238        | TTATCGACTCATCGGCGCTAAATGGGGGGATTTCAGCCTGTTAACATCAAGGCTGCAAGAC | 1297 |
|   | Qy | 1286        | CAGACACCCCATCTGCTCACCATCCAGCCGGACCTCAGCACCACCACCACCACCTACC    | 1345 |
|   | Db | 1298        | AAGACCTCTTGGCAGTGCCACCAGACCTCACTTCTGCTGCAGCCATGTACA           | 1348 |
|   | Qу | 1346        | AGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAA       | 1400 |
|   | Db | 1349        | GGGGGCCTGTGTATGCCTTGCATGATGTCTCTGATAAAATCCCAATGACCAATTCTCCGA  | 1408 |
|   | Qу | 1401        | TGGGCACCTGCTCAGCCCCCTGGGTGGCGCCGCCACACACTGCACCACACTCTC        | 1456 |
|   | Db | 1409        | TCCTGGACCCACTGCCCAATCTGAAGATTAAAGTTTATAACACCTCTGGAGCAGTCACCC  | 1468 |
|   | Qу | 1457        | CCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCC                    | 1498 |
|   | Db | 1469        | CCCAGGATGAACTCTCTGACTTCTCCTCCAAGCTGTCCCCACAGATTACCCAGTCTCTGT  | 1528 |

| Qy   | 1499 | AGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCA                          | 1534 |
|------|------|---|------|
| Db   | 1529 | TGGAGAATGAGACTCTGAACGTGAAGAACCAAAGCCTTGCACGGCAAACAGACCCATCCT  | 1588 |
| Qу   | 1535 | ACATGACCTATGGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTA  | 1594 |
| Db   | 1589 | GCACTGCATTTGGGACCTTCAACTCGTTAGGGGGCCACCTAGTAATTCCTAATTCAGGAG  | 1648 |
| Qy . | 1595 | TCAGCCTCCTCATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCA   | 1654 |
| Db . | 1649 | TGAGCTTGCTGATCCCAGCAGGGGCTGTTCCCCAAGGAAGAGTCTATGAAATGTATGT    | 1708 |
| Qу   | 1655 | CGCTGCACAAGCCGGAAGACGTGAGGTTGCCCCTAGCTGGCTG                   | 1714 |
| Db   | 1709 | CAGTCCACAGGAAGGAGGCATGAGACCACCTGTAGAAGACAGCCAGACGCTGCTGACAC   | 1768 |
| Qу   | 1715 | CCATCGTTAGCTGTGGACCCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGG  | 1774 |
| Db   | 1769 | CAGTGGTGAGCTGTGGCCCACCAGGAGCGCTGCTGACCCGACCCGTTGTGCTGACCATGC  | 1828 |
| Qу   | 1775 | ACCACTGTGGGGAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCG  | 1834 |
| Db   | 1829 | ACCACTGTGCTGAGCCCAACATGGATGACTGGCAGATCCAGCTCAAGCACCAGGCAGG    | 1888 |
| Qу   | 1835 | AGGGCAGCTGGGAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACT  | 1894 |
| Db   | 1889 | AGGGACCATGGGAGGATGTAGTGGTGGTCGGGGAGGAAAACTTCACCACTCCATGCTACA  | 1948 |
| Qу   | 1895 | GCCAGCTGGAGGCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGG  | 1954 |
| Db   | 1949 | TCCAGCTGGACCCAGAGGCCTGTCATATCCTGACGGAGACCCTCAGCACGTACGCCTTGG  | 2008 |
| Qу   | 1955 | TGGGAGAGGCCCTCAGCGTGCCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGG  | 2014 |
| Db   | 2009 | TGGGACAATCCATCACCAAAGCAGCCACACGTCTCAAATTGGCCATCTTTGGACCAC     | 2068 |
| Qy   | 2015 | TGGCCTGCACCTCCTCGAGTACAACATCCGGGTCTACTGCCTGC                  | 2074 |
| Db   | 2069 | TGTCCTGTTCCTCACTGGAGTACAGCATCCGCGTCTACTGCCTCGATGACACACAGGATG  | 2128 |
| Qy   | 2075 | CACTCAAGGAGGTGGTGCAGCTGGAGAAGCAGCTGGGGGGGACAGCTGATCCAGGAGCCAC | 2134 |
| Db   | 2129 | CCCTGAAGGAGGTCCTCCAGCTTGAGCGGCAGATGGGTGGG                     | 2188 |
| Qy   | 2135 | GGGTCCTGCACTTCAAGGACAGTTACCACAACCTGCGCCTATCCATCC              | 2194 |
| Db   | 2189 | AAACTTTGCATTTTAAGGGAAGTACCCACAACCTGCGCTTATCCATTCATGACATTGCCC  | 2248 |
| Qy   | 2195 | GCTCCCTGTGGAAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCT  | 2254 |
| Db   | 2249 | ACTCTCTCTGGAAGAGCAAACTGCCGGCTAAATACCAGGAGATTCCTTTTTACCACATCT  | 2308 |
| Qy   | 2255 | GGAATGCCACGCAGCGTACTTCCACCTTCACCCTGGAGCGTGTCAGCCCCAGCA        | 2314 |
| Db   | 2309 | GGAGTGGGTGCCAGAGGAACTTGCACTGCACCTTCACGCTGGAACGATTCAGTCTCAATA  | 2368 |

.

| Qу | 2315 | CTAGTGACCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCA | 2374 |
|----|------|--|------|
| Db | 2369 | CCCTGGAGCTCGTCTGCAAACTCTGTGTGCGGCAAGTCGAAGGAGAAGGGCAGATCTTCC | 2428 |
| Qу | 2375 | GCATCAACTTCAACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTG | 2434 |
| Db | 2429 | AGCTGAACTGCTCAGTATCAGAGGAACCCACTGGCATTGATTATCCCATCATGGATTCAG | 2488 |
| Qy | 2435 | AAGCGGGGGTCCCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGC | 2494 |
| Db | 2489 | CAGGCAGCATCACAACGATAGTTGGGCCCAACGCTTTCAGCATCCCCCTCCCAATAAGGC | 2548 |
| Qy | 2495 | AGAAGATAATTTCCAGCCTGGACCCACCCTGTAGGCGGGTGCCGACTGGCGGACTCTGG  | 2554 |
| Db | 2549 | AGAAGCTCTGCAGCAGCCTGGATGCACCCCAGACCCGGGGCCATGACTGGAGGATGCTGG | 2608 |
| Qу | 2555 | CCCAGAAACTCCACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAG | 2614 |
| Db | 2609 | CCCACAAGCTGAAATTGGACAGGTACCTAAATTATTTTGCTACGAAGTCGAGTCCCACTG | 2668 |
| Qy | 2615 | CCATGATCCTCAACCTGTGGGAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGC    | 2674 |
| Db | 2669 | GGGTGATCCTGGGAAGCCCAGAATTTCCCTGATGGCAACCTGAGCATGCTGG         | 2728 |
| Qу | 2675 | CTGCAGCAGTGGCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTG | 2734 |
| Db | 2729 | CAGCAGTGCTGGAAGAAATGGGACGACATGAAACCGTTGTTTCTTTGGCAGCAGAAGGAA | 2788 |
| Qу | 2735 | AGTGCTGAGGCCGGC 2749   |      |
| Db | 2789 | ATTACTGATGCAGCC 2803   |      |

Search completed: March 6, 2005, 05:24:52 Job time: 11708.8 secs

## GenCore version 5.1.6 Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2005, 14:42:51; Search time 1401 Seconds

(without alignments)

11628.216 Million cell updates/sec

Title: US-10-624-932-1

Perfect score: 2752

Sequence: 1 ccqcqqqqcccqqcccqq.....tqaqtqctqaqqccqqccaq 2752

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 segs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

1: geneseqn1980s:\*

2: geneseqn1990s:\*

3: geneseqn2000s:\*

4: geneseqn2001as:\*

5: geneseqn2001bs:\*

6: geneseqn2002as:\*

7: geneseqn2002bs:\*

8: geneseqn2003as:\*

9: geneseqn2003bs:\*

10: geneseqn2003cs:\*

11: geneseqn2003ds:\*

12: geneseqn2004as:\*

13: geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result<br>No. |        | Query<br>Match | Length | DB | ID       | Description        |
|---------------|--------|----------------|--------|----|----------|--------------------|
| 1             | · 2752 | 100.0          | 2752   | 6  | ABK37922 | Abk37922 cDNA enco |
| 2             | 2752   | 100.0          | 2752   | 12 | ADH71617 | Adh71617 Human gen |
| 3             | 2687.4 | 97.7           | 2697   | 6  | ABK52891 | Abk52891 Human net |
| 4             | 2676.4 | 97.3           | 2881   | 6  | ABK49422 | Abk49422 DNA encod |

```
2676.4
             97.3
                    2881
                          12
                              ADH71609
                                                        Adh71609 Human gen
    2674.8
             97.2
                    2881
                          12
                              ADH71649
                                                        Adh71649 Human gen
                                                        Adh71635 Human gen
 7
    2674.8
             97.2
                    2881
                          12
                              ADH71635
   2674.8
             97.2
                    2881
                          12
                             ADH71637
                                                        Adh71637 Human gen
    2674.8
             97.2
                    2881 12
                             ADH71641
                                                        Adh71641 Human gen
10
   2674.8
             97.2
                    2881
                          12
                              ADH71629
                                                        Adh71629 Human gen
    2674.8
             97.2
                    2881
                          12
                              ADH71631
                                                        Adh71631 Human gen
11
12
    2674.8
             97.2
                    2881
                          12
                              ADH71645
                                                        Adh71645 Human gen
13
   2674.8
             97.2
                    2881
                          12
                              ADH71627
                                                        Adh71627 Human gen
   2674.8
             97.2
                    2881
                         12
                             ADH71639
                                                        Adh71639 Human gen
14
                                                        Adh71643 Human gen
    2674.8
             97.2
                    2881
                         12
                             ADH71643
    2674.8
                    2881 12
                                                        Adh71625 Human gen
16
             97.2
                             ADH71625
    2674.8
                                                        Adh71647 Human gen
17
             97.2
                    2881 12
                             ADH71647
    2673.2
             97.1
                    2881
                              ADG42568
                                                        Adq42568 Novel hum
18
                          10
19
    2662.8
             96.8
                    2880
                          12 ADH71633
                                                        Adh71633 Human gen
                                                       Aak52261 Human pol
20
    2368.6
             86.1
                    2907
                          4 AAK52261
21
      2356
             85.6
                    3561 12 ADL06497
                                                       Adl06497 Human tum
22
      2356
             85.6
                    3580
                          6 ABK15169
                                                       Abk15169 Human REP
23
      2328
                    2575 12 ADH71621
                                                       Adh71621 Human gen
             84.6
                    2463 12 ADH71623
                                                       Adh71623 Human gen
24
   2266.2
             82.3
                                                     Aav52940 Rat UNC-5
25
      2259
             82.1
                    3014
                          2
                             AAV52940
                                                     Aas16843 Rat netri
26
    2252.2
             81.8
                    2697
                          6
                             AAS16843
                                                       Adn95100 Human LEC
27
    1635.4
             59.4
                    2635
                         11 ADN95100
                                                     Aav52941 Human UNC
Aac90958 Plasmid p
28
   1562.4
             56.8
                    1787 2 AAV52941
29
    1304.2
             47.4
                    9700 4
                             AAC90958
30
   1206.6
             43.8
                    1321 4
                             AAH9'9530
                                                      Aah99530 Human pro
31
    987.6
             35.9
                    1002 12 ADH71619
                                                       Adh71619 Human gen
     985.6
32
             35.8
                    1009
                         12 ADH71613
                                                       Adh71613 Human gen
       970
                    3646 5 AAS75738
                                                       Aas75738 DNA encod
33
             35.2
34
       970
             35.2
                    3646
                          13
                              ADR99257
                                                        Adr99257 Human unc
35
     968.8
             35.2
                    2796 10
                              AAL56266
                                                        Aal56266 Human thr
36
     943.4
             34.3
                    993 12 ADH71611
                                                       Adh71611 Human gen
     936.2
                    2860 6 ABT06279
37
             34.0
                                                      Abt06279 Human NOV
                                                     Abt06280 Human NOV
Abq93898 Human tra
38
     936.2
             34.0
                    2860 6
                             ABT06280
39
     913.6
             33.2
                    2895 6 ABQ93898
                                                      Adr07892 Full leng
Adr99249 Human 1RO
                    3933 13 ADR07892
40
     905.6
             32.9
                    2986 13 ADR99249
41
     904.2
             32.9
                                                     Aas21316 Human cDN
42
     902.4
             32.8
                    3884
                         4 AAS21316
43
     902.4
             32.8
                    3884 8 ACA03675
                                                      Aca03675 cDNA enco
44
     902.4
                    3884 8 ABX89213
                                                      Abx89213 DNA encod
             32.8
                    3884 8 ACD41867
     902.4
45
             32.8
                                                      Acd41867 Human sec
```

## ALIGNMENTS

```
RESULT 1
ABK37922
     ABK37922 standard; cDNA; 2752 BP.
XX
AC
    ABK37922;
XX
DT
     21-MAY-2002 (first entry)
XX
DE
     cDNA encoding Human protein NOV1.
XX
KW
```

Human; NOVX; ss; gene; cardiomyopathy; atherosclerosis; diabetes;

```
cell signal processing disorder; metabolic disorder; obesity; infection;
KW
     anorexia; cancer-associated cachexia; cancer; neurodegenerative disorder;
KW
     Alzheimer's disease; Parkinson's disease; immune disorder;
KW
     haematopoietic disorders; dyslipidaemia; pain; asthma; hypertension;
KW
     osteoporosis; Crohn's disease; multiple sclerosis; angina pectoris;
KW
     myocardial infarction; ulcer; allergy; benign prostatic hypertrophy;
KW
     psychosis; neurological disorder; anxiety; schizophrenia;
KW
KW
     manic depression; dementia; dyskinesia; Huntington's disease;
KW
     Gilles de la Tourette's syndrome; gene therapy.
XX
os
     Homo sapiens.
XX
PN
     WO200210216-A2.
XX
PD
     07-FEB-2002.
XX
PF
     30-JUL-2001; 2001WO-US024225.
XX
     28-JUL-2000; 2000US-0221409P.
PR
     04-AUG-2000; 2000US-0222840P.
PR
     04-AUG-2000; 2000US-0223752P.
PR
     04-AUG-2000; 2000US-0223762P.
PR
     04-AUG-2000; 2000US-0223769P.
PR
PR
     04-AUG-2000; 2000US-0223770P.
PR
     14-AUG-2000; 2000US-0225146P.
PR
     15-AUG-2000; 2000US-0225392P.
     15-AUG-2000; 2000US-0225470P.
PR
     16-AUG-2000; 2000US-0225697P.
PR
     01-FEB-2001; 2001US-0263662P.
PR
     05-APR-2001; 2001US-0281645P.
PR
XX
PΑ
     (CURA-) CURAGEN CORP.
XX
PΙ
     Padigaru M, Mezes P, Mishra V, Burgess C, Casman S, Grosse WM;
     Alsobrook JP, Lepley DM, Gerlach VL, Macdougall JR,
PΙ
XX
     WPI; 2002-180074/23.
DR
DR
     P-PSDB; AAU85403.
XX
PT
     New isolated cytoplasmic, nuclear, membrane bound, or secreted
PT
     polypeptide, useful for treating cardiomyopathy, atherosclerosis,
     infections, cancer, neurodegenerative, metabolic, hematopoietic and
PT
PT
     immune disorders.
XX
     Claim 9; Page 9-10; 213pp; English.
PS
XX
     The invention relates to an isolated cytoplasmic, nuclear, membrane
CC
     bound, or secreted polypeptide (NOVX, x=1-14) their variants or mature
CC
     form. Also included are the nucleic acids encoding the NOVX proteins, a
CC
     vector comprising the nucleic acid, a cell comprising the vector, an anti
CC
     -NOVX antibody and modulators of NOVX. NOVX, the nucleic acid and the
CC
CC
     antibody are useful for treating or preventing a NOVX-associated
CC
     disorder, where the disorder is selected from cardiomyopathy,
CC
     atherosclerosis, diabetes, a disorder related to cell signal processing
     and metabolic pathway modulation, metabolic disorders, obesity,
CC
CC
     infectious disease, anorexia, cancer-associated cachexia, cancer,
     neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,
CC
```

immune disorders, haematopoietic disorders, and the various CC dyslipidaemias, metabolic disturbances associated with obesity, the CC CC metabolic syndrome X and wasting disorders associated with chronic diseases, bacterial, fungal, protozoal and viral infections, pain, CC bulimia, asthma, hypertension, urinary retention, osteoporosis, Crohn's CC disease, multiple sclerosis, Albright Hereditary Osteodystrophy, angina CC pectoris, myocardial infarction, ulcer, allergy, benign prostatic CC CC hypertrophy, and psychotic and neurological disorders, including anxiety, CC schizophrenia, manic depression, delirium, dementia, and dyskinesias, such as Huntington's disease and Gilles de la Tourette's syndrome. The CC nucleic acid is useful in gene therapy. The present sequence encodes a CC CC NOVX protein XX

Sequence 2752 BP; 505 A; 937 C; 829 G; 481 T; 0 U; 0 Other;

. sq

Query Match 100.0%; Score 2752; DB 6; Length 2752;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2752; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| Qy | 1   | CCGCGGGGCCCGGCCCGCCCGCCTGCCCGCGCGCCATGGCCGTCCGGCCC 60            |
|----|-----|--|
| Db | . 1 |  |
| Qy | 61  | GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120 |
| Db | 61  | GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120 |
| Qy | 121 | CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180 |
| Db | 121 | CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180 |
| Qу | 181 | TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240 |
| Db | 181 | TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240 |
| Qў | 241 | AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300 |
| Db | 241 | AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300 |
| Qу | 301 | GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC 360 |
| Db | 301 |  |
| Qу | 361 | ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 420 |
| Db | 361 |  |
| Qу | 421 | TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 480 |
| Db | 421 |  |
| Qy | 481 | AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 540 |
| Db | 481 |  |
| Qу | 541 | ATCGTGCTGCCCTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 600      |

| Db   | 541  | ATCGTGCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG      | 600  |
|------|------|--|------|
| Qу   | 601  | AACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC     | 660  |
| Db   | 601  |  | 660  |
| Qy   | 661  | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC | 720  |
| Db   | 661  | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC | 720  |
| Qy   | 721  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG     | 780  |
| Db   | 721  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCAACGTGAACGGTGGGTG     | 780  |
| Qу   | 781  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG | 840  |
| Db . | 781  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG | 840  |
| Qу   | 841  | AGCCGGAGCTGCACCCACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT | 900  |
| Db   | 841  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT | 900  |
| Qу   | 901  | GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGGAGC | 960  |
| Db   | 901  | GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGGAGC | 960  |
| Qу   | 961  | AAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCA | 1020 |
| Db   | 961  | AAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCA | 1020 |
| Qу   | 1021 | GCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACC   | 1080 |
| Db   | 1021 | GCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACC   | 1080 |
| Qy   | 1081 | AGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTC | 1140 |
| Db   | 1081 | AGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTC | 1140 |
| Qу   | 1141 | ATCGCCGTGGCCGTCTGCTGCTGCTGCTGCTGCTCATCCTCGTTTATTGCCGG        | 1200 |
| Db   | 1141 | ATCGCCGTGGCCGTCTGCTGCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGG     | 1200 |
| QУ   | 1201 | AAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAG  | 1260 |
| Db   | 1201 | AAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAG  | 1260 |
| Qу   | 1261 | CCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGAC | 1320 |
| Db   | 1261 | CCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGAC | 1320 |
| Qy   | 1321 | CTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGC    | 1380 |
| Db   | 1321 | CTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGC    | 1380 |
| QУ   |      | CCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACA |      |
| Db   | 1381 | CCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACA | 1440 |

| Qу   | 1441 | CTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG   | 1500 |
|------|------|---|------|
| Db   | 1441 | CTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG  | 1500 |
| Qу   | 1501 | AACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTC  | 1560 |
| Db   | 1501 | AACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTC  | 1560 |
| QУ   | 1561 | CTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGATGCC   | 1620 |
| Db   | 1561 | CTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGATGCC  | 1620 |
| Qу   | 1621 | ATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGG  | 1680 |
| Db   | 1621 | ATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGG  | 1680 |
| Qу   | 1681 | TTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGC  | 1740 |
| Db   | 1681 | TTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGC  | 1740 |
| Qу   | 1741 | GTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGAC  | 1800 |
| Db   | 1741 | GTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGGAGCCCAGCCCTGAC | 1800 |
| Qy . | 1801 | AGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCTGCAC  | 1860 |
| Db . | 1801 | AGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCTGCAC  | 1860 |
| Qy   | 1861 | CTGGGCGAGGAGGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTAC    | 1920 |
| Db   | 1861 | CTGGGCGAGGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTAC    | 1920 |
| Qу   | 1921 | GTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCC  | 1980 |
| Db   | 1921 | GTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCC  | 1980 |
| Qу   | 1981 | GCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC           | 2040 |
| Db   | 1981 |   | 2040 |
| QУ   | 2041 | ATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAG  | 2100 |
| Db   | 2041 | ATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAG  | 2100 |
| Qу   | 2101 | AAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTAC  | 2160 |
| Db   | 2101 | AAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTAC  | 2160 |
| Qу   | 2161 | CACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTT  | 2220 |
| Db.  | 2161 | CACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTT  | 2220 |
| QУ   | 2221 | GTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCAC  | 2280 |
| Db   | 2221 | GTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCAC  | 2280 |
|      |      |   |      |

```
2281 TGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGG 2340
Qy
          2281 TGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGG 2340
Db
      2341 GTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGAC 2400
Qу
          2341 GTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGAC 2400
Db
      2401 ACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGC 2460
Qy
          2401 ACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGC 2460
Db
      2461 CCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCA 2520
Qу
          2461 CCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCA 2520
Db
      2521 CCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCAT 2580
Qy
          2521 CCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCAT 2580
Db
      2581 CTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGGGAGGCG 2640
Qу
          2581 CTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGGGAGGCG 2640
Db
      2641 CGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAG 2700
Qу
          Db
      2641 CGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAG 2700
      Qу
          Db
RESULT 2
ADH71617
   ADH71617 standard; DNA; 2752 BP.
ID
XX
AC
   ADH71617;
XX
DT
   25-MAR-2004
            (first entry)
XX
DE
   Human gene of the invention NOV21e SEQ ID NO:513.
XX
KW
   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
   vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
   obesity; diabetes; infectious disease; metabolic syndrome X;
KW
KW
   dyslipidaemia.
XX
os
   Homo sapiens.
XX
PN
   WO2003102155-A2.
XX
PD
   11-DEC-2003.
XX
PF
   03-JUN-2003; 2003WO-US017430.
XX
```

```
03-JUN-2002; 2002US-0385120P.
PR
     04-JUN-2002; 2002US-0385784P.
PR
     05-JUN-2002; 2002US-0386041P.
PR
PR
     05-JUN-2002; 2002US-0386047P.
     06-JUN-2002; 2002US-0386376P.
PR
     06-JUN-2002; 2002US-0386453P.
PR
     06-JUN-2002; 2002US-0386864P.
PR
PR
     06-JUN-2002; 2002US-0387016P.
     07-JUN-2002; 2002US-0386796P.
PR
PR
     07-JUN-2002; 2002US-0386816P.
     07-JUN-2002; 2002US-0386931P.
PR
PR
     07-JUN-2002; 2002US-0386942P.
     07-JUN-2002; 2002US-0386971P.
PR
     07-JUN-2002; 2002US-0387262P.
PR
     08-JUN-2002; 2002US-0296960P.
PR
PR
     10-JUN-2002; 2002US-0387400P.
PR
     10-JUN-2002; 2002US-0387535P.
     11-JUN-2002; 2002US-0387610P.
PR
PR
     11-JUN-2002; 2002US-0387625P.
     11-JUN-2002; 2002US-0387634P.
PR
     11-JUN-2002; 2002US-0387668P.
PR
     11-JUN-2002; 2002US-0387696P.
PR
PR
     11-JUN-2002; 2002US-0387702P.
PR
     11-JUN-2002; 2002US-0387836P.
     11-JUN-2002; 2002US-0387859P.
PR
PR
     12-JUN-2002; 2002US-0387933P.
PR
     12-JUN-2002; 2002US-0387934P.
PR
     12-JUN-2002; 2002US-0387960P.
PR
     12-JUN-2002; 2002US-0388022P.
     12-JUN-2002; 2002US-0388096P.
PR
PR
     13-JUN-2002; 2002US-0389123P.
PR
     14-JUN-2002; 2002US-0389118P.
PR
     14-JUN-2002; 2002US-0389120P.
PR
     14-JUN-2002; 2002US-0389144P.
     14-JUN-2002; 2002US-0389146P.
PR
     17-JUN-2002; 2002US-0389729P.
PR
     17-JUN-2002; 2002US-0389742P.
PR
     18-JUN-2002; 2002US-0389884P.
PR
PR
     19-JUN-2002; 2002US-0390006P.
ΡR
     19-JUN-2002; 2002US-0390209P.
PR
     21-JUN-2002; 2002US-0390763P.
PR
     17-JUL-2002; 2002US-0396706P.
     06-AUG-2002; 2002US-0401628P.
PR
     09-AUG-2002; 2002US-0402156P.
PR
     09-AUG-2002; 2002US-0402256P.
PR
PR
     09-AUG-2002; 2002US-0402389P.
PR
     12-AUG-2002; 2002US-0402786P.
PR
     12-AUG-2002; 2002US-0402816P.
PR
     12-AUG-2002; 2002US-0402821P.
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
PR
     13-AUG-2002; 2002US-0403459P.
PR
     13-AUG-2002; 2002US-0403531P.
PR
     13-AUG-2002; 2002US-0403532P.
PR
PR
     13-AUG-2002; 2002US-0403563P.
PR
     13-AUG-2002; 2002US-0406317P.
PR
     15-AUG-2002; 2002US-0403617P.
```

```
PR
     26-AUG-2002; 2002US-0406182P.
     26-AUG-2002; 2002US-0406355P.
PR
PR
     27-AUG-2002; 2002US-0406240P.
PR
     12-SEP-2002; 2002US-0410084P.
     20-SEP-2002; 2002US-0412528P.
PR
PR
     23-SEP-2002; 2002US-0412731P.
     30-SEP-2002; 2002US-0414801P.
PR
     30-SEP-2002; 2002US-0414839P.
PR
     30-SEP-2002; 2002US-0414840P.
PR
PR
     30-SEP-2002; 2002US-0414954P.
     09-OCT-2002; 2002US-0417186P.
PR
PR
     09-OCT-2002; 2002US-0417406P.
PR
     23-OCT-2002; 2002US-0420639P.
     28-OCT-2002; 2002US-0421156P.
PR
     31-OCT-2002; 2002US-0422690P.
PR
     01-NOV-2002; 2002US-0423130P.
PR
     05-NOV-2002; 2002US-00423798.
PR
PR
     05-NOV-2002; 2002US-0423798P.
PR
     12-NOV-2002; 2002US-0425453P.
XX
PA
     (CURA-) CURAGEN CORP.
XX
PΙ
    Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PΙ
     Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI
     Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PΙ
     Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
ΡI
     Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
ΡI
     Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D,
                                                                 Rastelli L;
PΙ
     Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
ΡI
     Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PΙ
     Zhong H;
XX
DR
    WPI; 2004-081935/08.
DR
     P-PSDB; ADH71618.
XX
PT
    New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT
     treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
     obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS
     Example 21; SEQ ID NO 513; 1880pp; English.
XX
CC
     The invention relates to a novel isolated polypeptide (NOVX). A
CC .
    polypeptide of the invention has cytostatic, immunomodulator,
CC
    neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
     antilipaemic activity, and may have a use in gene therapy, and as a
CC
     vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC
     any of the 303 fully defined nucleotide sequences given in the
CC
     specification. The polypeptide is useful in the manufacture of a
CC
    medicament for treating a syndrome associated with a human disease. The
CC
    polypeptide, polynucleotide and antibody are useful in diagnosing,
CC
     treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC
    Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC
    diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC
     further used as hybridisation probes, in chromosome mapping, tissue
CC
     typing, preventive medicine, and pharmacogenomics. The present sequence
```

encodes a NOVX polypeptide of the invention.

CC

XX

100.0%; Score 2752; DB 12; Length 2752; Query Match 100.0%; Pred. No. 0: Best Local Similarity Matches 2752; Conservative Mismatches 0; 0; 0; Indels 0; Gaps 1 CCGCGGGGCCCGCCCGGCCGCCCGCCGCCGCGCGGCCATGGCCGTCCGGCCC 60 Qy 1 CCGCGGGGCCCGGCCCGCCCGCCCGCCGCCGCGCGCCATGGCCGTCCGGCCC 60 Db 61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTTGGCTCCGCGGCTCGGGTGCC 120 Qy \_61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120 Db 121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180 Qу 121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180 Db 181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240 Qу 181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240 Db 241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300 Qу 241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300 Db 301 GACCACGTGATCGAGCGCAGCAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC 360 Qy 301 GACCACGTGATCGAGCGCAGCACGGGGGGGGGGGGGCTGCCCACCATGGAGGTCCGC 360 Db 361 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 420 Qу ``` 361 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 420 Db 421 TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 480 Qу 421 TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 480 Db 481 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 540 Qy 481 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 540 Db 541 ATCGTGCTGCCTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 600 Qy 541 ATCGTGCTGCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 600 Db 601 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 660 Qу 601 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 660 Db 661 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 720 Qу 661 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 720 Db Qу 

Db

|   | QУ   | 781  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG         | 840  |
|---|------|------|--|------|
|   | Db   | 781  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG         | 840  |
|   | Qу   | 841  | ${\tt AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT}$ | 900  |
|   | Db   | 841  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT         | 900  |
|   | Qу   | 901  | GTCCAGAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGGAGC          | 960  |
|   | Db   | 901  | GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGGAGC         | 960  |
|   | Qу   | 961  | AAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCA         | 1020 |
| * | Db   | 961  | AAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGACCCA         | 1020 |
|   | Qу   | 1021 | GCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACC           | 1080 |
|   | Db   | 1021 | GCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGTACC           | 1080 |
|   | QУ   | 1081 | AGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTC         | 1140 |
|   | Db   | 1081 | AGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGCCTC         | 1140 |
|   | Qу   | 1141 | ATCGCCGTGGCCGTCTGCTGGTCCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGCCGG          | 1200 |
|   | Db   | 1141 | ATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTCATCCTCGTTTATTGCCGG               | 1200 |
|   | Qу   | 1201 | AAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAG          | 1260 |
|   | Db   | 1201 | AAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTCCAG         | 1260 |
| · | Qу   | 1261 | CCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGAC         | 1320 |
|   | Db   | 1261 | CCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCGGAC         | 1320 |
|   | Qу   | 1321 | CTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGC            | 1380 |
|   | Db   | 1321 | CTCAGCACCACCACCACCACCAGGGGCAGTCTCTGTCCCCGGCAGGATGGGCCCAGC            | 1380 |
|   | Qy   | 1381 | CCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACA         | 1440 |
|   | Db   | 1381 | CCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCACACA         | 1440 |
|   | Qу   | 1441 | CTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG          | 1500 |
|   | Db . | 1441 | CTGCACCACGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACCCAG           | 1500 |
|   | Qу   | 1501 | AACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTC         | 1560 |
|   | Db   | 1501 | AACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAACTTC         | 1560 |
|   | QУ   | 1561 | CTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGATGCC          | 1620 |
|   | Db   | 1561 | CTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGATGCC         | 1620 |

.

| Qу   | 1621 | ATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGG | 1680 |
|------|------|--|------|
| Db   | 1621 | ATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTGAGG | 1680 |
| QУ   | 1681 | TTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGC | 1740 |
| Db   | 1681 | TTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCTGGC | 1740 |
| Qу   | 1741 | GTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGAC | 1800 |
| Db   | 1741 | GTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCTGAC | 1800 |
| Qу   | 1801 | AGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCTGCAC | 1860 |
| Db   | 1801 | AGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTGCTGCAC | 1860 |
| QУ   | 1861 | CTGGGCGAGGAGGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTAC   | 1920 |
| Db   | 1861 | CTGGGCGAGGAGGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCCTGCTAC   | 1920 |
| Qу   | 1921 | GTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCC | 1980 |
| Db   | 1921 | GTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTGGCTGCC | 1980 |
| Qу   | 1981 | GCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC          | 2040 |
| Db   | 1981 | GCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC          | 2040 |
| Qу   | 2041 | ATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAG | 2100 |
| Db   | 2041 | ATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAGCTGGAG | 2100 |
| Qу   | 2101 | AAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTAC | 2160 |
| Db   | 2101 | AAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGACAGTTAC | 2160 |
| Qу   | 2161 | CACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTT | 2220 |
| Db   | 2161 | CACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAGCTCCTT | 2220 |
| Qу   | 2221 | GTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCAC | 2280 |
| Db . | 2221 | GTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTACTTGCAC | 2280 |
| Qу   | 2281 | TGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGG | 2340 |
| Db   | 2281 | TGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAGCTGTGG | 2340 |
| Qу   | 2341 | GTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGAC | 2400 |
| Dp,  | 2341 | GTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACCAAGGAC | 2400 |
| Qу   | 2401 | ACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGC | 2460 |
| Db   | 2401 | ACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTGGTGGGC | 2460 |
| Ov   | 2461 | CCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCA | 2520 |

.

```
2461 CCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTGGACCCA 2520
Db
       2521 CCCTGTAGGCGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCAT 2580
Qy
          2521 CCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGACAGCCAT 2580
Db
       2581 CTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGGGAGGCG 2640
Qy
          2581 CTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGGGAGGCG 2640
Db
       2641 CGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAG 2700
Qy
          2641 CGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCCAG 2700
Db
       Qу
          Db
RESULT 3
ABK52891
   ABK52891 standard; DNA; 2697 BP.
XX
AC
   ABK52891;
XX
DT
   27-AUG-2002 (first entry)
XX
DE
   Human netrin binding membrane receptor UNC5H-1 DNA sequence #1.
XX
   Netrin binding membrane receptor; receptor; UNC5H-1; gene; ds; human;
KW
KW
   nootropic; neuroprotective; cytostatic; antiparkinsonian;
   cerebroprotective; cancer; central nervous system; CNS; stroke;
KW
   Parkinson's disease; multiple sclerosis; Alzheimer's disease.
KW
XX
os
   Homo sapiens.
XX
FH
   Kev
                Location/Qualifiers
                1. .2697
FT
   CDS
                /*tag= a
FT
FT
                /product= "Netrin binding membrane receptor UNC5H-1"
XX
   WO200233080-A2.
PN
XX
PD
   25-APR-2002.
XX
PF
   15-OCT-2001; 2001WO-EP011891.
XX
PR
   16-OCT-2000; 2000US-0240061P.
XX
    (FARB ) BAYER AG.
PA
XX ·
ΡI
   Koehler RH;
XX
DR
   WPI; 2002-463314/49.
DR
   P-PSDB; AAU97899.
XX
```

Novel human netrin binding membrane receptor polypeptide and polynucleotides for identifying modulating agents useful in treating diseases e.g. Parkinson's disease, multiple sclerosis, stroke, Alzheimer's disease.

XX PS

Claim 1; Fig 1; 94pp; English.

CC CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC XX

SO

PT

PT

PT PT

> This invention relates to the DNA and protein sequences of a novel purified human netrin binding membrane receptor, UNC5H-1. The DNA sequence of the invention is useful as a probe for detecting a nucleic acid encoding the UNC5H-1 protein in a biological sample. The sequences of the invention are useful to screen for agents which decrease the activity of the UNC5H-1 protein. The sequences are also useful for screening agents which regulate (modulate) the activity of the protein of the invention. A pharmaceutical composition containing the protein of the invention or a reagent that modulates the activity of the UNC5H-1 protein may be useful for treating a UNC5H-1 dysfunction related disease such as cancer or a central nervous system (CNS) disorders (e.g, Parkinson's disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion proteins comprising the UNC5H-1 protein are useful for generating antibodies and for in various assay systems, and the protein can be used as a bait protein in a two-hybrid assay or three-hybrid assay. The method of the invention is useful for detecting a coding sequence for the UNC5H-1 protein. The present sequence represents a DNA sequence encoding the human netrin binding membrane receptor UNC5H-1 protein of the invention

Sequence 2697 BP; 503 A; 906 C; 807 G; 481 T; 0 U; 0 Other;

Query Match 97.7%; Score 2687.4; DB 6; Length 2697; Best Local Similarity 99.8%; Pred. No. 0; Matches 2691; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

| Matches | 209. | i; conservative | 0;      | MISMacches   | 0,      | Inders    | 0,     | Gaps     | 0,  |
|---------|------|-----------------|---------|--------------|---------|-----------|--------|----------|-----|
| Qy      | 46   | ATGGCCGTCCGGCCC |         |              | •       |           |        |          | 105 |
| Db      | 1    | ATGGCCGTCCGGCCC |         |              |         |           |        |          | 60  |
| Qу      | 106  | CGCGGCTCGGGTGCC |         |              |         |           |        |          | 165 |
| Db      | 61   | CGCGGCTCGGGTGCC |         |              |         |           |        | <i>.</i> | 120 |
| Qу      | 166  | GACCTGCTTCCCCAC |         |              |         |           |        |          | 225 |
| Db      | 121  | GACCTGCTTCCCCAC |         |              |         |           |        |          | 180 |
| Qу      | 226  | GTGCTGCTTGTGTGC |         |              |         |           |        |          | 285 |
| Db      | 181  | GTGCTGCTTGTGTGC |         |              |         |           |        |          | 240 |
| Qу      | 286  | TGGGTGCGCCAGGTG |         |              |         |           |        |          | 345 |
| Db      | 241  | TGGGTGCGCCAGGTG |         |              |         |           |        |          | 300 |
| Qу      | 346  | ACCATGGAGGTCCGC |         |              |         |           |        |          | 405 |
| Db      | 301  | ACCATGGAGGTCCGC |         |              |         |           |        |          | 360 |
| Qу      | 406  | GAATACTGGTGCCAG | TGCGTGG | CATGGAGCTCC' | TCGGGCA | .CCACCAAG | AGTCAG | AAGGCC   | 465 |

| Db | 361  | GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC  | 420  |
|----|------|---|------|
| Qy | 466  | TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG  | 525  |
| Db | 421  | TACATCCGCATAGCCTATTTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG  | 480  |
| Qу | 526  | TCCCTGGAGCAGGCATCGTGCTGCCCTGCCGTCCACCGGAGGCCATCCCTCCAGCCGAG   | 585  |
| Db | 481  | TCCCTGGAGCAGGCATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG   | 540  |
| Qу | 586  | GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC  | 645  |
| Db | 541  |   | 600  |
| Qy | 646  | ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC  | 705  |
| Db | 601  |   | 660  |
| Qу | 706  | TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC  | 765  |
| Db | 661  | TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTAC     | 720  |
| Qу | 766  | GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCC      | 825  |
| Db | 721  |   | 780  |
| Qу | 826  | GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC  | 885  |
| Db | 781  |   | 840  |
| Qу | 886  | TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC  | 945  |
| Db | 841  | TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTGGACGGCAGC  | 900  |
| Qу | 946  | TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT  | 1005 |
| Db | 901  | •   | 960  |
| Qy | 1006 | GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGGGAGTGCCAGGGCACTGACCTGGAC | 1065 |
| Db | 961  | GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGAC   | 1020 |
| Qy | 1066 | ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC  | 1125 |
| Db | 1021 | ACCCGCAACTGTACCAGTGACCTCTGTGTACACACTGCTTCTGGCCCTGAGGACGTGGCC  | 1080 |
| Qу | 1126 | CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTGTCCTCATC   | 1185 |
| Db | 1081 |   | 1140 |
| Qу | 1186 | CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC  | 1245 |
| Db | 1141 |   | 1200 |
| Qy | 1246 | ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC  | 1305 |

| Dp .    | 1201 | ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC | 1260 |
|---------|------|--|------|
| Qy      | 1306 | ACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG    | 1365 |
| Db      | 1261 | ACCATCCAGCCGGACCTCAGCACCACCACCACCAGGGCAGTCTCTGTCCCCGG        | 1320 |
| Qу      | 1366 | CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT | 1425 |
| Db      | 1321 | CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT | 1380 |
| Qу      | 1426 | GGCGGCCACACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC    | 1485 |
| Db      | 1381 | GGCGGCCACACACTGCACCACACTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCC      | 1440 |
| Qy      | 1486 | CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT | 1545 |
| Db      | 1441 | CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT | 1500 |
| Qy      | 1546 | GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC | 1605 |
| Db      | 1501 | GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGAATCAGCCTCCTC | 1560 |
| Qy      | 1606 | ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG  | 1665 |
| Db      | 1561 | ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG  | 1620 |
| Qу      | 1666 | CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC | 1725 |
| Db      | 1621 | CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC | 1680 |
| Qу      | 1726 | TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG  | 1785 |
| Db      | 1681 | TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG  | 1740 |
| Qу      | 1786 | GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG | 1845 |
| Db      | 1741 | GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG | 1800 |
| Qy      | 1846 | GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG | 1905 |
| Db      | 1801 | GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG | 1860 |
| Qy      | 1906 | GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC | 1965 |
| Db      | 1861 | GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC | 1920 |
| Qy<br>· | 1966 | CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC  | 2025 |
| Db      | 1921 | CTCAGCGTGGCTGCCCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC  | 1980 |
| Qy      | 2026 | TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG | 2085 |
| Db      | 1981 | TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG | 2040 |
| QУ      | 2086 | GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC | 2145 |
| Db ·    | 2041 | GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC | 2100 |

```
2101 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2160
Db
      2206 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2265
Qу
         2161 AAGAGTAAGCTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACG 2220
Db
      2266 CAGCGGTACTTGCACTGCACCTTGACCCTGGAGCGTCAGCCCCAGCACTAGTGACCTG 2325
Qy
         2221 CAGCGGTACTTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTG 2280
Db
      2326 GCCTGCAAGCTGTGGGTGGGAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2385
Qy
         2281 GCCTGCAAGCTGTGGGTGGGAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTC 2340
Db
      2386 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2445
Qу
         2341 AACATCACCAAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTC 2400
Db
      2446 CCAGCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2505
Qу
         2401 CCAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATT 2460
Db
      2506 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2565
Qy
         2461 TCCAGCCTGGACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTC 2520
Db
      2566 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2625
Qу
         2521 CACCTGGACAGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTC 2580
Db
      Qу
         Db
      2686 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2742
Qy
         2641 GCTGGACTGGGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGA 2697
Db
         Ţ
RESULT 4
ABK49422
ID
   ABK49422 standard; DNA; 2881 BP.
XX
AC
   ABK49422;
XX
DT
   15-JUL-2002
            (first entry)
XX
   DNA encoding human UNC5-like protein NOV1.
DE
XX
KW
   Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;
   cell signal processing; metabolic pathway modulation; cancerous tissue;
KW
KW
   antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;
KW
   chromosome 13; gene; ds.
XX
```

2146 TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG 2205

Qу

```
Homo sapiens.
os
XX
                  Location/Oualifiers
FH
    Key
                  87. .2786
FT
    CDS
FT
                  /*tag= a
FT
                  /product= "Human UNC5-like protein NOV1"
XX
PN
    WO200229038-A2.
XX
PD
    11-APR-2002.
XX
PF
    04-OCT-2001; 2001WO-US031377.
XX
    04-OCT-2000; 2000US-0237862P.
PR
XX
PA
    (CURA-) CURAGEN CORP.
XX
PΙ
    Herrmann JL,
                Rastelli L,
                           Shimkets RA;
XX
DR
    WPI; 2002-340104/37.
    P-PSDB; AAU79939.
DR
XX
PT
    Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for
PT
    treating cardiomyopathy, artherosclerosis, and cancer.
XX
    Claim 8; Page 7-8; 180pp; English.
PS
XX
CC
    The present invention relates to a new NOVX polypeptide having a 900
CC
    (NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or 331 (NOV6)
    residue amino acid sequence, as given in the specification. The novel
CC
    polypeptide, and its encoding polynucleotide, are used to treat
CC
CC
    cardiomyopathy, atherosclerosis, cancer or a disease related to cell
    signal processing and metabolic pathway modulation, in a human. Detecting
CC
CC
    the polypeptide or polynucleotide is useful for identifying cancerous
    tissue. The antibody can be used to treat diabetes or cancer. The host
CC
    cells can be used to produce non-human transgenic animals useful in drug
CC
    screening. The present nucleic acid sequence is that of the human UNC5-
CC
    like NOV1 gene located on chromosome 13. This sequence encodes the human
CC
    UNC5-like protein NOV1 of the invention
CC
XX
    Sequence 2881 BP; 526 A; 985 C; 868 G; 502 T; 0 U; 0 Other;
SO
                             Score 2676.4; DB 6;
 Query Match
                      97.3%;
                                                Length 2881;
 Best Local Similarity
                      98.98;
                             Pred. No. 0;
                            0; Mismatches 21; Indels
 Matches 2728; Conservative
                                                                   3;
          Qу
           Db
         61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120
Qу
            102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161
Db
        121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180
Qу
            162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221
Db
```

| Qу   | 181  | TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC       | 240  |
|------|------|--|------|
| Db   | 222  | TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC       | 281  |
| Qy   | 241  | AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCCCAGGTG        | 300  |
| Db · | 282  | AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG       | 341  |
| Qy   | 301  | GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC       | 360  |
| Db   | 342  | GACCACGTGATCGAGCGCACCACGACCACGGGGGGCAGCAGTGGTGAGCCGACCATGGAGGTCCGC | 401  |
| Qy   | 361  | ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG       | 420  |
| Db   | 402  | ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG       | 461  |
| Qy . | 421  | TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC       | 480  |
| Db   | 462  | TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC       | 521  |
| Qу   | 481  | AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC       | 540  |
| Db   | 522  | AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC       | 581  |
| Qу   | 541  | ATCGTGCTGCCCTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG            | 600  |
| Db   | 582  | ATCGTGCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG            | 641  |
| Qу   | 601  | AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC       | 660  |
| Db   | 642  | AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC       | 701  |
| Qу   | 661  | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC       | 720  |
| Db   | 702  | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGCCCAACTACACCTGCGTGGCCAAGAAC       | 761  |
| Qу   | 721  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG           | 780  |
| Db   | 762  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG           | 821  |
| Qу   | 781  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG       | 840  |
| Db   | 822  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG       | 881  |
| Qу   | 841  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT       | 900  |
| Db   | 882  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT       | 941  |
| Qy   | 901  | GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG          | 957  |
| Db   | 942  | GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG       | 1001 |
| Qy   | 958  | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC       | 1017 |
| Db   | 1002 | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC       | 1061 |

| Qу  | 1018 | CCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT   | 1077 |
|-----|------|--|------|
| Db  | 1062 |  | 1121 |
| Qу  | 1078 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC | 1137 |
| Db  | 1122 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC | 1181 |
| Qу  | 1138 | $\tt CTCATCGCCGTGGCCTGCTGCTGCTGCTGCTGCTGCTCATCCTCGTTTATTGC$  | 1197 |
| Db  | 1182 |  | 1241 |
| Qу  | 1198 | CGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC  | 1257 |
| Db  | 1242 |  | 1301 |
| Qy  | 1258 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG | 1317 |
| Db  | 1302 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG | 1361 |
| Qy  | 1318 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC    | 1377 |
| Db  | 1362 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC    | 1418 |
| Qy  | 1378 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC | 1437 |
| Db  | 1419 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC | 1478 |
| Qу  | 1438 | ACACTGCACCACGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC   |      |
| Db  | 1479 | ACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC |      |
| Qу  | 1498 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC | 1557 |
| Db  | 1539 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC | 1598 |
| Qy. | 1558 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT  | 1617 |
| Db  | 1599 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT  | 1658 |
| Qу  | 1618 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG | 1677 |
| Db  | 1659 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG | 1718 |
| Qу  | 1678 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT | 1737 |
| Db  | 1719 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT | 1778 |
| Qy  | 1738 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1797 |
| Db  | 1779 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1838 |
| Qу  | 1798 | GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTG    | 1854 |
| Db  | 1839 | GACAGCTGGAGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG       | 1898 |
| Ov  | 1855 | CTGCACCTGGGCGAGGGGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC | 1914 |

•

| Db | 1899 | CTGCACCTGGGCGAGGGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC        | 19,58 |
|----|------|--|-------|
| Qу | 1915 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG       | 1974  |
| Db | 1959 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG       | 2018  |
| Qу | 1975 | GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC          | 2034  |
| Db | 2019 | GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC          | 2078  |
| Qy | 2035 | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG       | 2094  |
| Db | 2079 | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG       | 2138  |
| Qу | 2095 | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC       | 2154  |
| Db | 2139 | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC       | 2198  |
| Qy | 2155 | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG       | 2214  |
| Db | 2199 | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG       | 2258  |
| Qу | 2215 | CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC       | 2274  |
| Db | 2259 | $\tt CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC$ | 2318  |
| QУ | 2275 | TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG       | 2334  |
| Db | 2319 | TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG       | 2378  |
| Qу | 2335 | CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC       | 2394  |
| Db | 2379 | $\tt CTGTGGGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG$                       | 2438  |
| Qу |      | AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG       |       |
| Db |      | AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG       |       |
| Qу |      | GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG       |       |
| Db | 2499 | GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG       | 2558  |
| Qу |      | GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC       |       |
| Db | 2559 | GACCCACCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC        | 2618  |
| Qу |      | AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGG        |       |
| Db | 2619 | AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGG        | 2678  |
| QУ |      | GAGGCGCGCACTTCCCCAACGCCAACCTCAGCCAGCTGGCTG                         |       |
| Db |      | GAGGCGCGCACTTCCCCAACGCCAACCTCAGCCAGCTGGCTG                         |       |
| Qу | 2695 | GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCC           | 752   |
|    |      |  |       |

PR

12-JUN-2002; 2002US-0388096P.

```
RESULT 5
ADH71609
     ADH71609 standard; DNA; 2881 BP.
ID .
XX
AC
     ADH71609;
XX
DT
     25-MAR-2004
                  (first entry)
XX
DE
     Human gene of the invention NOV21a SEQ ID NO:505.
XX
KW
     ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW
     anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
     vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
     obesity; diabetes; infectious disease; metabolic syndrome X;
KW
     dyslipidaemia.
XX
OS
     Homo sapiens.
XX
PN
     WO2003102155-A2.
XX
PD
     11-DEC-2003.
XX
     03-JUN-2003; 2003WO-US017430.
PF
XX
PR
     03-JUN-2002; 2002US-0385120P.
PR
     04-JUN-2002; 2002US-0385784P.
     05-JUN-2002; 2002US-0386041P.
PR
PR
     05-JUN-2002; 2002US-0386047P.
PR
     06-JUN-2002; 2002US-0386376P.
PR
     06-JUN-2002; 2002US-0386453P.
PR
     06-JUN-2002; 2002US-0386864P.
PR
     06-JUN-2002; 2002US-0387016P.
     07-JUN-2002; 2002US-0386796P.
PR
     07-JUN-2002; 2002US-0386816P.
PR
     07-JUN-2002; 2002US-0386931P.
PR
PR
     07-JUN-2002; 2002US-0386942P.
PR
     07-JUN-2002; 2002US-0386971P.
     07-JUN-2002; 2002US-0387262P.
PR
PR
     08-JUN-2002; 2002US-0296960P.
     10-JUN-2002; 2002US-0387400P.
PR
     10-JUN-2002; 2002US-0387535P.
PR
     11-JUN-2002; 2002US-0387610P.
PR
     11-JUN-2002; 2002US-0387625P.
PR
PR
     11-JUN-2002; 2002US-0387634P.
PR
     11-JUN-2002; 2002US-0387668P.
PR
     11-JUN-2002; 2002US-0387696P.
PR
     11-JUN-2002; 2002US-0387702P.
     11-JUN-2002; 2002US-0387836P.
PR
     11-JUN-2002; 2002US-0387859P.
PR
     12-JUN-2002; 2002US-0387933P.
PR
PR
     12-JUN-2002; 2002US-0387934P.
PR
     12-JUN-2002; 2002US-0387960P.
PR
     12-JUN-2002; 2002US-0388022P.
```

```
13-JUN-2002; 2002US-0389123P.
PR
     14-JUN-2002; 2002US-0389118P.
PR
     14-JUN-2002; 2002US-0389120P.
PR
PR
     14-JUN-2002; 2002US-0389144P.
PR
     14-JUN-2002; 2002US-0389146P.
PR
     17-JUN-2002; 2002US-0389729P.
     17-JUN-2002; 2002US-0389742P.
PR
     18-JUN-2002; 2002US-0389884P.
PR
     19-JUN-2002; 2002US-0390006P.
PR
PR
     19-JUN-2002; 2002US-0390209P.
     21-JUN-2002; 2002US-0390763P.
PR
     17-JUL-2002; 2002US-0396706P.
PR
PR
     06-AUG-2002; 2002US-0401628P.
     09-AUG-2002; 2002US-0402156P.
PR
     09-AUG-2002; 2002US-0402256P.
PR
     09-AUG-2002; 2002US-0402389P.
PR
     12-AUG-2002; 2002US-0402786P.
PR
     12-AUG-2002; 2002US-0402816P.
PR
     12-AUG-2002; 2002US-0402821P.
PR
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
PR
     13-AUG-2002; 2002US-0403459P.
PR
     13-AUG-2002; 2002US-0403531P.
PR
     13-AUG-2002; 2002US-0403532P.
PR
     13-AUG-2002; 2002US-0403563P.
PR
     13-AUG-2002; 2002US-0406317P.
PR
PR
     15-AUG-2002; 2002US-0403617P.
PR
     26-AUG-2002; 2002US-0406182P.
     26-AUG-2002; 2002US-0406355P.
PR
     27-AUG-2002; 2002US-0406240P.
PR
     12-SEP-2002; 2002US-0410084P.
PR
     20-SEP-2002; 2002US-0412528P.
PR
     23-SEP-2002; 2002US-0412731P.
PR
     30-SEP-2002; 2002US-0414801P.
PR
PR
     30-SEP-2002; 2002US-0414839P.
     30-SEP-2002; 2002US-0414840P.
PR
     30-SEP-2002; 2002US-0414954P.
PR
     09-OCT-2002; 2002US-0417186P.
PR
PR
     09-OCT-2002; 2002US-0417406P.
     23-OCT-2002; 2002US-0420639P.
PR
     28-OCT-2002; 2002US-0421156P.
PR
     31-OCT-2002; 2002US-0422690P.
PR
PR
     01-NOV-2002; 2002US-0423130P.
     05-NOV-2002; 2002US-00423798.
PR
     05-NOV-2002; 2002US-0423798P.
PR
PR
     12-NOV-2002; 2002US-0425453P.
XX
PA
     (CURA-) CURAGEN CORP.
XX
PI
     Alsobrook JP, Alvarez E,
                                Anderson DW, Boldog FL, Casman SJ;
     Catterton E, Chapoval A,
PI
                                Crabtree-Bokor JR, Edinger SR, Ellerman K;
     Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PΙ
     Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PΙ
PI
    Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PI
                  Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
     Padigaru M,
ΡI
     Rieger DK, Rothenberg ME,
                                 Sciore P, Shenoy SG,
                                                        Shimkets RA;
                  Spytek KA, Stone DJ, Vernet CAM, Voss EZ,
     Smithson G,
PΙ
```

```
PΙ
    Zhong H;
XX
    WPI: 2004-081935/08.
DR
    P-PSDB; ADH71610.
DR
XX
    New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT
    treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
    obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
PT
XX
PS
    Example 21; SEQ ID NO 505; 1880pp; English.
XX
    The invention relates to a novel isolated polypeptide (NOVX). A
CC
CC
    polypeptide of the invention has cytostatic, immunomodulator,
    neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
    antilipaemic activity, and may have a use in gene therapy, and as a
CC
    vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC
    any of the 303 fully defined nucleotide sequences given in the
CC
    specification. The polypeptide is useful in the manufacture of a
CC
    medicament for treating a syndrome associated with a human disease. The
CC
    polypeptide, polynucleotide and antibody are useful in diagnosing,
CC
    treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC
    Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC
    diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC
    further used as hybridisation probes, in chromosome mapping, tissue
CC
    typing, preventive medicine, and pharmacogenomics. The present sequence
CC
    encodes a NOVX polypeptide of the invention.
CC
XX
    Sequence 2881 BP; 526 A; 985 C; 868 G; 502 T; 0 U; 0 Other;
SO
                       97.3%;
                             Score 2676.4; DB 12; Length 2881;
 Best Local Similarity
                       98.9%;
                            Pred. No. 0;
 Matches 2728; Conservative
                             0; Mismatches
                                            21; Indels
                                                         9; Gaps
          1 CCGCGGGGCCCGGCCCGCCCGCCGCCGCGCGCCATGGCCGTCCGGCCC 60
Qу
            42 CCGCGGGGCCCGGCCCGGCCCGCCCGCCGCGCCGGGCCATGGCCGTCCGGCCC 101
Db
         61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120
Qy
            102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161
Db
         121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180
Qy
            162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221
Db
         181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240
Qу
            222 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 281
Db
         241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300
QУ
            282 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 341
Db
         301 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC 360
Qу
            111 111111111111
Db
         342 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGC 401
```

| Qy | 361  | ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG  | 420  |
|----|------|---|------|
| Db | 402  | ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG  | 461  |
| Qу | 421  | TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC  | 480  |
| Db | 462  | TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC  | 521  |
| Qу | 481  | AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC  | 540  |
| Db | 522  | AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC  | 581  |
| QУ | 541  | ATCGTGCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG       | 600  |
| Db | 582  | ATCGTGCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG       | 641  |
| Qу | 601  | AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC  | 660  |
| Db | 642  | AACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC      | 701  |
| Qу | 661  | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC  | 720  |
| Db | 702  | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC  | 761  |
| Qy | 721  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTG   | 780  |
| Db | 762  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG      | 821  |
| Qy | 781  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGGCGGCTGGCAGAAACGG | 840  |
| Db | 822  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG  | 881  |
| Qу | 841  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT  | 900  |
| Db | 882  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT  | 941  |
| Qу | 901  | GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG     | 957  |
| Db | 942  | GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG  | 1001 |
| Qy | 958  | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC  | 1017 |
| Db | 1002 | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC  | 1061 |
| Qу | 1018 | CCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT    | 1077 |
| Db | 1062 | CCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT    | 1121 |
| Qy | 1078 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC  | 1137 |
| Db | 1122 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC  | 1181 |
| Qy | 1138 | CTCATCGCCGTGGCCGTCTGCTGGTCCTGCTGCTGCTCCTCATCCTCGTTTATTGC      | 1197 |
| Db | 1182 | CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTGCTCCTC                 | 1241 |
| Qу | 1198 | CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC  | 1257 |

| Db | 1242 | CGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC  | 1301 |
|----|------|--|------|
| Qу | 1258 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG | 1317 |
| Db | 1302 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG | 1361 |
| Qу | 1318 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC    | 1377 |
| Db | 1362 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC    | 1418 |
| Qy | 1378 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC | 1437 |
| Db | 1419 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC | 1478 |
| Qу | 1438 | ACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC | 1497 |
| Db | 1479 | ACACTGCACCACACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC          | 1538 |
| Qу | 1498 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC | 1557 |
| Db | 1539 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC | 1598 |
| Qу | 1558 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT  | 1617 |
| Db | 1599 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT  | 1658 |
| Qу | 1618 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG | 1677 |
| Db | 1659 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG | 1718 |
| Qу | 1678 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT | 1737 |
| Db | 1719 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT | 1778 |
| Qу | 1738 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1797 |
| Db | 1779 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1838 |
| Qу | 1798 | GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTG    | 1854 |
| Db | 1839 | GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG | 1898 |
| Qу |      | CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC |      |
| Db | 1899 | CTGCACCTGGGCGAGGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC   | 1958 |
| Qу | 1915 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG | 1974 |
| Db | 1959 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTŢTGCCCTGGTGGGAGAGGCCCTCAGCGTG | 2018 |
| Qу | 1975 | GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC    | 2034 |
| Db |      | GCTGCCGCCAAGCCTCCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC     |      |
| Qу | 2035 | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG | 2094 |

. . .

٠

```
Db
     2079 TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG 2138
     2095 CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC 2154
Qy
        2139 CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC 2198
Db
     2155 AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG 2214
Qу
        2199 AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG 2258
Db
     2215 CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC 2274
Qy
        2259 CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC 2318
Db
     2275 TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG 2334
Qу
        2319 TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG 2378
Db
     2335 CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC 2394
Qу
        2379 CTGTGGGTGTGGCAGGTGGAGGCGACGGGCAGAGCTTCAGCATCAACATCACC 2438
Db
     2395 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2454
Qу
        2439 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2498
Db
     2455 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2514
Qу
        2499 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2558
Db
     2515 GACCCACCTGTAGGCGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2574
Qу
        2559 GACCCACCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2618
Db
     2575 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGG 2634
Qy
        2619 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGG 2678
Db
     Qу
        Db
     Qу
        Db
RESULT 6
ID
   ADH71649 standard; DNA; 2881 BP.
XX
```

```
ADH71649
ID ADH71649 standard; DNA; 2881 BP.
XX
AC ADH71649;
XX
DT 25-MAR-2004 (first entry)
XX
DE Human gene of the invention NOV21u SEQ ID NO:545.
```

XX

```
KW
     ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW
     anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
     vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
     obesity; diabetes; infectious disease; metabolic syndrome X;
KW
     dyslipidaemia.
XX
os
     Homo sapiens.
XX
PN
     WO2003102155-A2.
XX
PD
     11-DEC-2003.
XX
PF
     03-JUN-2003; 2003WO-US017430.
XX
     03-JUN-2002; 2002US-0385120P.
PR
PR
     04-JUN-2002; 2002US-0385784P.
     05-JUN-2002; 2002US-0386041P.
PR
PR
     05-JUN-2002; 2002US-0386047P.
     06-JUN-2002; 2002US-0386376P.
PR
     06-JUN-2002; 2002US-0386453P.
PR
     06-JUN-2002; 2002US-0386864P.
PR
     06-JUN-2002; 2002US-0387016P.
PR
     07-JUN-2002; 2002US-0386796P.
PR
     07-JUN-2002; 2002US-0386816P.
PR
     07-JUN-2002; 2002US-0386931P.
PR
     07-JUN-2002; 2002US-0386942P.
PR
PR
     07-JUN-2002; 2002US-0386971P.
PR
     07-JUN-2002; 2002US-0387262P.
PR
     08-JUN-2002; 2002US-0296960P.
     10-JUN-2002; 2002US-0387400P.
PR
     10-JUN-2002; 2002US-0387535P.
PR
     11-JUN-2002; 2002US-0387610P.
PR
     11-JUN-2002; 2002US-0387625P.
PR
PR
     11-JUN-2002; 2002US-0387634P.
PR
     11-JUN-2002; 2002US-0387668P.
PR
     11-JUN-2002; 2002US-0387696P.
     11-JUN-2002; 2002US-0387702P.
PR
     11-JUN-2002; 2002US-0387836P.
PR
PR
     11-JUN-2002; 2002US-0387859P.
PR
     12-JUN-2002; 2002US-0387933P.
PR
     12-JUN-2002; 2002US-0387934P.
     12-JUN-2002; 2002US-0387960P.
PR
     12-JUN-2002; 2002US-0388022P.
PR
PR
     12-JUN-2002; 2002US-0388096P.
     13-JUN-2002; 2002US-0389123P.
PR
     14-JUN-2002; 2002US-0389118P.
PR
     14-JUN-2002; 2002US-0389120P.
PR
PR
     14-JUN-2002; 2002US-0389144P.
PR
     14-JUN-2002; 2002US-0389146P.
PR
     17-JUN-2002; 2002US-0389729P.
PR
     17-JUN-2002; 2002US-0389742P.
     18-JUN-2002; 2002US-0389884P.
PR
PR
     19-JUN-2002; 2002US-0390006P.
PR
     19-JUN-2002; 2002US-0390209P.
PR
     21-JUN-2002; 2002US-0390763P.
PR
     17-JUL-2002; 2002US-0396706P.
```

PR

06-AUG-2002; 2002US-0401628P.

```
09-AUG-2002; 2002US-0402156P.
PR
     09-AUG-2002; 2002US-0402256P.
PR
     09-AUG-2002; 2002US-0402389P.
PR
     12-AUG-2002; 2002US-0402786P.
PR
     12-AUG-2002; 2002US-0402816P.
PR
     12-AUG-2002; 2002US-0402821P.
PR
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
PR
     13-AUG-2002; 2002US-0403459P.
PR
PR
     13-AUG-2002; 2002US-0403531P.
PR
     13-AUG-2002; 2002US-0403532P.
PR
     13-AUG-2002; 2002US-0403563P.
     13-AUG-2002; 2002US-0406317P.
PR
     15-AUG-2002; 2002US-0403617P.
PR
     26-AUG-2002; 2002US-0406182P.
PR
     26-AUG-2002; 2002US-0406355P.
PR
     27-AUG-2002; 2002US-0406240P.
PR
PR
     12-SEP-2002; 2002US-0410084P.
PR
     20-SEP-2002; 2002US-0412528P.
     23-SEP-2002; 2002US-0412731P.
PR
     30-SEP-2002; 2002US-0414801P.
PR
     30-SEP-2002; 2002US-0414839P.
PR
     30-SEP-2002; 2002US-0414840P.
PR
     30-SEP-2002; 2002US-0414954P.
PR
PR
     09-OCT-2002; 2002US-0417186P.
PR
     09-OCT-2002; 2002US-0417406P.
     23-OCT-2002; 2002US-0420639P.
PR
     28-OCT-2002; 2002US-0421156P.
PR
     31-OCT-2002; 2002US-0422690P.
PR
     01-NOV-2002; 2002US-0423130P.
PR
     05-NOV-2002; 2002US-00423798.
PR
PR
     05-NOV-2002; 2002US-0423798P.
PR
     12-NOV-2002; 2002US-0425453P.
XX
PA
     (CURA-) CURAGEN CORP.
XX
     Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PΙ
PI
     Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
     Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PΙ
     Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PΙ
PΙ
     Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PΙ
     Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D,
                                                                 Rastelli L;
PΙ
     Rieger DK, Rothenberg ME,
                                 Sciore P, Shenoy SG, Shimkets RA;
PΙ
     Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ,
PI
     Zhong H;
XX
DR
     WPI; 2004-081935/08.
DR
     P-PSDB; ADH71650.
XX
PT
     New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT
     treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
     obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
     Example 21; SEQ ID NO 545; 1880pp; English.
PS
XX
CC
     The invention relates to a novel isolated polypeptide (NOVX). A
CC
     polypeptide of the invention has cytostatic, immunomodulator,
```

neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and CC antilipaemic activity, and may have a use in gene therapy, and as a CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising CC any of the 303 fully defined nucleotide sequences given in the CC specification. The polypeptide is useful in the manufacture of a CC medicament for treating a syndrome associated with a human disease. The CC CC polypeptide, polynucleotide and antibody are useful in diagnosing, CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia, CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are CC further used as hybridisation probes, in chromosome mapping, tissue CC CC typing, preventive medicine, and pharmacogenomics. The present sequence CC encodes a NOVX polypeptide of the invention. XX

SQ Sequence 2881 BP; 526 A; 986 C; 868 G; 501 T; 0 U; 0 Other;

Query Match 97.2%; Score 2674.8; DB 12; Length 2881; Best Local Similarity 98.9%; Pred. No. 0; Matches 2727; Conservative 0; Mismatches 22; Indels 9; Gaps 3,

| 114001 |     | ,, 551251745175 5, 1125111251125 22, 1111121 5, 545          | _   |
|--------|-----|--|-----|
| Qу     | 1   | CCGCGGGGCCCGCCCGCCCGCCCGCCCGCCCGCCCATGGCCGTCCGGCCC           | 60  |
| Db     | 42  | CCGCGGGGCCCGCCCGCCCGCCCGCCCGCCCGCCCCCCC                      | 101 |
| Qy     | 61  | GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC | 120 |
| Db     | 102 | GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC | 161 |
| Qy     | 121 | CAGCAGAGTGCCACCGTGGCCAACCCGGACCTGCTTCCCCAC                   | 180 |
| Db     | 162 | CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGCTGCCAACCCGGACCTGCTTCCCCAC | 221 |
| Qy     | 181 | TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC | 240 |
| Db     | 222 | TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC | 281 |
| Qу     | 241 | AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG | 300 |
| Db     | 282 | AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG | 341 |
| Qy     | 301 | GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC | 360 |
| Db     | 342 | GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGC | 401 |
| Qy     | 361 | ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG | 420 |
| Db     | 402 | ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG | 461 |
| Qу     | 421 | TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC | 480 |
| Db     | 462 | TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC | 521 |
| Qy     | 481 | AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC | 540 |
| Db     | 522 | $ \begin{array}{cccccccccccccccccccccccccccccccccccc$        | 581 |
| Qу     | 541 | ATCGTGCTGCCCTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG      | 600 |
|        |     |  |     |

| Db   | 582  | ATCGTGCTGCCCTCCACCGGAGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG        | 641  |
|------|------|---|------|
| Qу   | 601  | AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC  | 660  |
| Db   | 642  | AACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC      | 701  |
| Qу   | 661  | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC  | 720  |
| Db   | 702  | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC  | 761  |
| Qy . | 721  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG      | 780  |
| Db   | 762  | ATCGTGGCACGTCGCCGCAGCGCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG       | 821  |
| Qу   | 781  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG  | 840  |
| Db   | 822  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG  | 881  |
| Qу   | 841  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT  | 900  |
| Db   | 882  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT  | 941  |
| QУ   | 901  | GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG     | 957  |
| Db   | 942  | GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG  | 1001 |
| Qу   | 958  | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC  | 1017 |
| Db   | 1002 | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC  | 1061 |
| Qу   | 1018 | CCAGCACCCGCAACGGAGGGGAGGGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT | 1077 |
| Db   | 1062 | CCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT    | 1121 |
| QΥ . | 1078 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC  | 1137 |
| Db   | 1122 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC  | 1181 |
| QУ   | 1138 | CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTGCTCCTC                 | 1197 |
| Db   | 1182 | CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTCTCATCCTCGTTTATTGC      | 1241 |
| Qу   | 1198 | CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC  | 1257 |
| Db   | 1242 | CGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC   | 1301 |
| Qу   | 1258 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG  | 1317 |
| Db   | 1302 | CAGCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG   | 1361 |
| Qу   | 1318 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC     | 1377 |
| Db . | 1362 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC     | 1418 |
| Qу   | 1378 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC  | 1437 |

| Db   | 1419 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC | 1478 |
|------|------|--|------|
| Qу   | 1438 | ACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC | 1497 |
| Db   | 1479 | ACACTGCACCACACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC          | 1538 |
| Qу   | 1498 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC | 1557 |
| Db   | 1539 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC | 1598 |
| Qy   | 1558 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT  | 1617 |
| Db   | 1599 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT  | 1658 |
| Qy   | 1618 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG | 1677 |
| Db   | 1659 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG | 1718 |
| Qу   | 1678 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT | 1737 |
| Db   | 1719 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT | 1778 |
| Qу   | 1738 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1797 |
| Db   | 1779 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1838 |
| δÀ   | 1798 | GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTG    | 1854 |
| Db   | 1839 | GACAGCTGGAGCCTCCAAAAAGCAGTCGTGCGAGGCAGCTGGGAGCAGGATGTG       | 1898 |
| Qy . | 1855 | CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC | 1914 |
| Db   | 1899 | CTGCACCTGGGCGAGGAGGCCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC  | 1958 |
| Qy   | 1915 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG | 1974 |
| Db   | 1959 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG | 2018 |
| Qy   | 1975 | GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC    |      |
| Db   | 2019 | GCTGCCGCCAAGCCTCCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC     |      |
| Qy   |      | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG | 2094 |
| Db   |      | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG | 2138 |
| Qy   | 2095 | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC | 2154 |
| Db   | 2139 | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC | 2198 |
| Qy   | 2155 | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG | 2214 |
| Db   | 2199 | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG | 2258 |
| Qу   | 2215 | CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC | 2274 |
| Db   | 2259 | CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC | 2318 |

```
2275 TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG 2334
Qy
         Db
      2319 TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG 2378
      2335 CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC 2394
Qy
         2379 CTGTGGGTGTGGCAGGTGGAGGGCGACGGCCAGAGCTTCAGCATCAACTTCAACATCACC 2438
Db
      2395 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2454
Qу
         2439 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2498
Db
      2455 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2514
Qу
         2499 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2558
Db
      2515 GACCCACCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2574
Qу
         2559 GACCCACCTGTAGGCGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2618
Db
      2575 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGG 2634
Qy
         2619 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGCGG 2678
Db
      Qy
         Db
      Qy
         Db
RESULT 7
ADH71635
   ADH71635 standard; DNA; 2881 BP.
XX
AC
   ADH71635;
XX
DT
   25-MAR-2004
           (first entry)
XX
DE
   Human gene of the invention NOV21n SEQ ID NO:531.
XX
KW
   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW
   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
   vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
   obesity; diabetes; infectious disease; metabolic syndrome X;
KW
   dyslipidaemia.
XX
os
   Homo sapiens.
XX
   WO2003102155-A2.
PN
XX
PD
   11-DEC-2003.
XX
PF
   03-JUN-2003; 2003WO-US017430.
```

```
XX
     03-JUN-2002; 2002US-0385120P.
PR
     04-JUN-2002; 2002US-0385784P.
PR
PR
     05-JUN-2002; 2002US-0386041P.
PR
     05-JUN-2002; 2002US-0386047P.
     06-JUN-2002; 2002US-0386376P.
PR
     06-JUN-2002; 2002US-0386453P.
PR
PR
     06-JUN-2002; 2002US-0386864P.
     06-JUN-2002; 2002US-0387016P.
PR
PR
     07-JUN-2002; 2002US-0386796P.
     07-JUN-2002; 2002US-0386816P.
PR
PR
     07-JUN-2002; 2002US-0386931P.
     07-JUN-2002; 2002US-0386942P.
PR
     07-JUN-2002; 2002US-0386971P.
PR
     07-JUN-2002; 2002US-0387262P.
PR
     08-JUN-2002; 2002US-0296960P.
PR
PR
     10-JUN-2002; 2002US-0387400P.
     10-JUN-2002; 2002US-0387535P.
PR
PR
     11-JUN-2002; 2002US-0387610P.
     11-JUN-2002; 2002US-0387625P.
PR
     11-JUN-2002; 2002US-0387634P.
PR
     11-JUN-2002; 2002US-0387668P.
PR
PR
     11-JUN-2002; 2002US-0387696P.
PR
     11-JUN-2002; 2002US-0387702P.
PR
     11-JUN-2002; 2002US-0387836P.
     11-JUN-2002; 2002US-0387859P.
PR
     12-JUN-2002; 2002US-0387933P.
PR
     12-JUN-2002; 2002US-0387934P.
PR
     12-JUN-2002; 2002US-0387960P.
PR
     12-JUN-2002; 2002US-0388022P.
PR
     12-JUN-2002; 2002US-0388096P.
PR
PR
     13-JUN-2002; 2002US-0389123P.
     14-JUN-2002; 2002US-0389118P.
PR
PR
     14-JUN-2002; 2002US-0389120P.
     14-JUN-2002; 2002US-0389144P.
PR
     14-JUN-2002; 2002US-0389146P.
PR
     17-JUN-2002; 2002US-0389729P.
PR
     17-JUN-2002; 2002US-0389742P.
PR
     18-JUN-2002; 2002US-0389884P.
PR
     19-JUN-2002; 2002US-0390006P.
PR
PR
     19-JUN-2002; 2002US-0390209P.
PR
     21-JUN-2002; 2002US-0390763P.
     17-JUL-2002; 2002US-0396706P.
PR
     06-AUG-2002; 2002US-0401628P.
PR
     09-AUG-2002; 2002US-0402156P.
PR
PR
     09-AUG-2002; 2002US-0402256P.
PR
     09-AUG-2002; 2002US-0402389P.
PR
     12-AUG-2002; 2002US-0402786P.
PR
     12-AUG-2002; 2002US-0402816P.
PR
     12-AUG-2002; 2002US-0402821P.
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
PR
PR
     13-AUG-2002; 2002US-0403459P.
PR
     13-AUG-2002; 2002US-0403531P.
PR
     13-AUG-2002; 2002US-0403532P.
PR
     13-AUG-2002; 2002US-0403563P.
PR
     13-AUG-2002; 2002US-0406317P.
```

```
PR . 15-AUG-2002; 2002US-0403617P.
     26-AUG-2002; 2002US-0406182P.
PR
PR
     26-AUG-2002; 2002US-0406355P.
     27-AUG-2002; 2002US-0406240P.
PR
     12-SEP-2002; 2002US-0410084P.
PR
     20-SEP-2002; 2002US-0412528P.
PR
     23-SEP-2002; 2002US-0412731P.
PR
     30-SEP-2002; 2002US-0414801P.
PR
PR
     30-SEP-2002; 2002US-0414839P.
PR
     30-SEP-2002; 2002US-0414840P.
     30-SEP-2002; 2002US-0414954P.
PR
     09-OCT-2002; 2002US-0417186P.
PR
     09-OCT-2002; 2002US-0417406P.
PR
     23-OCT-2002; 2002US-0420639P.
PR
     28-OCT-2002; 2002US-0421156P.
PR
     31-OCT-2002; 2002US-0422690P.
PR
     01-NOV-2002; 2002US-0423130P.
PR
     05-NOV-2002; 2002US-00423798.
PR
     05-NOV-2002; 2002US-0423798P.
PR
PR
     12-NOV-2002; 2002US-0425453P.
XX
PA
     (CURA-) CURAGEN CORP.
XX
PΙ
    Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
     Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PΙ
                   Gangolli EA, Gerlach VL, Gorman L, Gunther E,
PI
     Ettenberg S,
PI
     Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PΙ
    Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PΙ
     Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
     Rieger DK, Rothenberg ME,
                                 Sciore P, Shenoy SG, Shimkets RA;
PΙ
PΙ
     Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PΙ
     Zhong H;
XX
DR
    WPI; 2004-081935/08.
DR
     P-PSDB; ADH71636.
XX
     New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT
     treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
PT
     obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS
     Example 21; SEQ ID NO 531; 1880pp; English.
XX
CC
     The invention relates to a novel isolated polypeptide (NOVX). A
CC
     polypeptide of the invention has cytostatic, immunomodulator,
     neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
CC
     antilipaemic activity, and may have a use in gene therapy, and as a
     vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC
     any of the 303 fully defined nucleotide sequences given in the
CC
CC
     specification. The polypeptide is useful in the manufacture of a
CC
     medicament for treating a syndrome associated with a human disease. The
CC
     polypeptide, polynucleotide and antibody are useful in diagnosing,
CC
     treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC
     Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
     diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC
CC
     further used as hybridisation probes, in chromosome mapping, tissue
     typing, preventive medicine, and pharmacogenomics. The present sequence
CC
CC
     encodes a NOVX polypeptide of the invention.
```

Query Match

Sequence 2881 BP; 527 A; 985 C; 867 G; 502 T; 0 U; 0 Other;

Score 2674.8; DB 12; Length 2881;

97.2%;

Best Local Similarity 98.9%; Pred. No. 0; 0; Matches 2727; Conservative Mismatches 22; Indels 1 CCGCGGGGCCCGGCCCGCCCGCCCGCCCGCCCGCCCATGGCCGTCCGGCCC 60 Qy Db 42 CCGCGGGCCCCGCCCGCCCGCCCGCCCGCCCGCGCCATGGCCGTCCGGCCC 101 61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120 Qy . 102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161 Db 121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180 Qу 162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221 Db 181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240 Qy 222 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 281 Db 241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300 Qy 282 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 341 Db 301 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC 360 Qy 342 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGC 401 Db 361 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 420 Qу 402 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 461 Db 421 TGCGTGGCATGGAGCTCCTCGGGCACCACAAGAGTCAGAAGGCCTACATCCGCATAGCC 480 Qу 462 TGCGTGGCATGGAGCTCCTCGGGCACCACAAGAGTCAGAAGGCCTACATCCGCATAGCC 521 Db 481 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 540 Qу 522 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 581 Db 541 ATCGTGCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 600 Qу 582 ATCGTGCTGCCTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 641 Db 601 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 660 Qу 642 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 701 Db 661 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 720 Qу 702 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 761 Db Qy 

|   | Db   | 762  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG     | 821  |
|---|------|------|--|------|
|   | Qy   | 781  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG | 840  |
|   | Db   | 822  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG | 881  |
|   | Qy   | 841  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT | 900  |
|   | Db   | 882  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT | 941  |
|   | Qy   | 901  | GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG    | 957  |
|   | Db   | 942  | GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG | 1001 |
| • | Qy   | 958  | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC | 1017 |
|   | Db   | 1002 |  | 1061 |
|   | Qy   | 1018 | CCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT   | 1077 |
|   | Db   | 1062 |  | 1121 |
|   | Qy   | 1078 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC | 1137 |
|   | Db   | 1122 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC | 1181 |
|   | Qy . | 1138 | CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTCATCCTCGTTTATTGC       | 1197 |
|   | Db   | 1182 | CTCATCGCCGTGGCCTGCTGCTGCTGCTGCTGCTCCTCATCCTCGTTTATTGC        | 1241 |
|   | Qу   | 1198 | CGGAAGAAGGAGGGGCTGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC  | 1257 |
|   | Db   | 1242 | CGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC  | 1301 |
|   | Qу   | 1258 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG | 1317 |
|   | Db   | 1302 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG | 1361 |
|   | Qу   | 1318 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC    | 1377 |
|   | Db   | 1362 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC    | 1418 |
|   | Qу   | 1378 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC | 1437 |
|   | Db   | 1419 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC | 1478 |
|   | Qу   | 1438 | ACACTGCACCACGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC   | 1497 |
|   | Db   | 1479 | ACACTGCACCACACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC          | 1538 |
|   | Qу   | 1498 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC | 1557 |
|   | Db   | 1539 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC | 1598 |
|   | Qу   | 1558 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT  | 1617 |
|   | Db   | 1599 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGAT | 1658 |

| Qу   | 1618 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG  | 1677 |
|------|------|---|------|
| Db   | 1659 |   | 1718 |
| Qу   | 1678 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT  | 1737 |
| Db   | 1719 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT  | 1778 |
| Qy   | 1738 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT  | 1797 |
| Db   | 1779 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT  | 1838 |
| Qу   | 1798 | GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTG     | 1854 |
| Db   | 1839 | GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG  | 1898 |
| Qу   | 1855 | CTGCACCTGGGCGAGGAGGCCCCCCCCCCCCCTCTACTACTGCCAGCTGGAGGCCAGTGCC | 1914 |
| Db   | 1899 | CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC  | 1958 |
| Qy   | 1915 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG  | 1974 |
| Db   | 1959 | TGCTACATCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG  | 2018 |
| Qy   | 1975 | GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC     | 2034 |
| Db   | 2019 | GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC     | 2078 |
| Qу   | 2035 | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG  | 2094 |
| Db   | 2079 | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG  | 2138 |
| Qу   | 2095 | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC  | 2154 |
| Db   | 2139 | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC  | 2198 |
| Qy   | 2155 | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG  | 2214 |
| Db   | 2199 | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG  | 2258 |
| Qy . | 2215 | CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC  | 2274 |
| Db . | 2259 | CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC  | 2318 |
| Qу   | 2275 | TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG  | 2334 |
| Db   | 2319 | TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG  | 2378 |
| Qу   | 2335 | CTGTGGGTGTGGCAGGTGGAGGĞCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC  | 2394 |
| Db   | 2379 |   | 2438 |
| Qу   | 2395 | AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG  | 2454 |
| Db   | 2439 |   | 2498 |

.

.

```
2455 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2514
Qy
          2499 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2558
Db
      2515 GACCCACCCTGTAGGCGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2574
Qγ
          2559 GACCCACCTGTAGGCGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2618
Db
      2575 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGG 2634
Qу
          2619 AGCCATCTCAGCTTCTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGG 2678
Db
      Qу
          Db
      Qу
          Db
RESULT 8
ADH71637
   ADH71637 standard; DNA; 2881 BP.
XX
AC
   ADH71637;
XX
DT
   25-MAR-2004
             (first entry)
XX
DΕ
   Human gene of the invention NOV210 SEQ ID NO:533.
XX
KW
   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW
   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
   vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
KW
   obesity; diabetes; infectious disease; metabolic syndrome X;
KW
   dyslipidaemia.
XX
os
   Homo sapiens.
XX
PN
   WO2003102155-A2.
XX
PD
   11-DEC-2003.
XX
PF
   03-JUN-2003; 2003WO-US017430.
XX
   03-JUN-2002; 2002US-0385120P.
PR
   04-JUN-2002; 2002US-0385784P.
PR
   05-JUN-2002; 2002US-0386041P.
PR
   05-JUN-2002; 2002US-0386047P.
PR
   06-JUN-2002; 2002US-0386376P.
PR
PR
   06-JUN-2002; 2002US-0386453P.
   06-JUN-2002; 2002US-0386864P.
PR
   06-JUN-2002; 2002US-0387016P.
PR
   07-JUN-2002; 2002US-0386796P.
PR
   07-JUN-2002; 2002US-0386816P.
PR
   07-JUN-2002; 2002US-0386931P.
PR
PR
   07-JUN-2002; 2002US-0386942P.
```

```
07-JUN-2002; 2002US-0386971P.
PR
PR
     07-JUN-2002; 2002US-0387262P.
PR
     08-JUN-2002; 2002US-0296960P.
     10-JUN-2002; 2002US-0387400P.
PR
PR
     10-JUN-2002; 2002US-0387535P.
PR
     11-JUN-2002; 2002US-0387610P.
     11-JUN-2002; 2002US-0387625P.
PR
PR
     11-JUN-2002; 2002US-0387634P.
PR
     11-JUN-2002; 2002US-0387668P.
PR
     11-JUN-2002; 2002US-0387696P.
     11-JUN-2002; 2002US-0387702P.
PR
     11-JUN-2002; 2002US-0387836P.
PR
     11-JUN-2002; 2002US-0387859P.
PR
     12-JUN-2002; 2002US-0387933P.
PR
     12-JUN-2002; 2002US-0387934P.
PR
     12-JUN-2002; 2002US-0387960P.
PR
     12-JUN-2002; 2002US-0388022P.
PR
     12-JUN-2002; 2002US-0388096P.
PR
     13-JUN-2002; 2002US-0389123P.
PR
     14-JUN-2002; 2002US-0389118P.
PR
     14-JUN-2002; 2002US-0389120P.
PR
     14-JUN-2002; 2002US-0389144P.
PR
     14-JUN-2002; 2002US-0389146P.
PR
PR
     17-JUN-2002; 2002US-0389729P.
PR
     17-JUN-2002; 2002US-0389742P.
PR
     18-JUN-2002; 2002US-0389884P.
     19-JUN-2002; 2002US-0390006P.
PR
     19-JUN-2002; 2002US-0390209P.
PR
     21-JUN-2002; 2002US-0390763P.
PR
     17-JUL-2002; 2002US-0396706P.
PR
     06-AUG-2002; 2002US-0401628P.
PR
     09-AUG-2002; 2002US-0402156P.
PR
     09-AUG-2002; 2002US-0402256P.
PR
PR
     09-AUG-2002; 2002US-0402389P.
     12-AUG-2002; 2002US-0402786P.
PR
     12-AUG-2002; 2002US-0402816P.
PR
     12-AUG-2002; 2002US-0402821P.
PR
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
PR
     13-AUG-2002; 2002US-0403459P.
PR
     13-AUG-2002; 2002US-0403531P.
PR
PR
     13-AUG-2002; 2002US-0403532P.
     13-AUG-2002; 2002US-0403563P.
PR
     13-AUG-2002; 2002US-0406317P.
PR
     15-AUG-2002; 2002US-0403617P.
PR
     26-AUG-2002; 2002US-0406182P.
PR
PR
     26-AUG-2002; 2002US-0406355P.
PR
     27-AUG-2002; 2002US-0406240P.
PR
     12-SEP-2002; 2002US-0410084P.
     20-SEP-2002; 2002US-0412528P.
PR
     23-SEP-2002; 2002US-0412731P.
PR
     30-SEP-2002; 2002US-0414801P.
PR
PR
     30-SEP-2002; 2002US-0414839P.
     30-SEP-2002; 2002US-0414840P.
PR
PR
     30-SEP-2002; 2002US-0414954P.
     09-OCT-2002; 2002US-0417186P.
PR
PR
     09-OCT-2002; 2002US-0417406P.
```

```
23-OCT-2002; 2002US-0420639P.
PR
    28-OCT-2002; 2002US-0421156P.
PR
    31-OCT-2002; 2002US-0422690P.
PR
    01-NOV-2002; 2002US-0423130P.
PR
PR
    05-NOV-2002; 2002US-00423798.
PR
    05-NOV-2002; 2002US-0423798P.
    12-NOV-2002; 2002US-0425453P.
PR
XX
PA
    (CURA-) CURAGEN CORP.
XX
PI
    Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
    Catterton E, Chapoval A, Crabtree-Bokor JR,
                                                Edinger SR, Ellerman K;
ΡI
    Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PΙ
    Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI
    Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PΙ
PΙ
    Padigaru M, Patturajan M,
                              Pena CEA, Peyman JA, Raha D, Rastelli L;
PI
    Rieger DK, Rothenberg ME,
                               Sciore P, Shenoy SG, Shimkets RA;
PI
    Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PΙ
    Zhong H;
XX
    WPI; 2004-081935/08.
DR
DR
    P-PSDB; ADH71638.
XX
PT
    New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT
    treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
    obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS
    Example 21; SEQ ID NO 533; 1880pp; English.
XX
CC
    The invention relates to a novel isolated polypeptide (NOVX). A
    polypeptide of the invention has cytostatic, immunomodulator,
CC
    neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
CC
    antilipaemic activity, and may have a use in gene therapy, and as a
CC
    vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
    any of the 303 fully defined nucleotide sequences given in the
CC
    specification. The polypeptide is useful in the manufacture of a
CC
CC
    medicament for treating a syndrome associated with a human disease. The
CC
    polypeptide, polynucleotide and antibody are useful in diagnosing,
CC
    treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
    Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC
CC
    diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC
    further used as hybridisation probes, in chromosome mapping, tissue
    typing, preventive medicine, and pharmacogenomics. The present sequence
CC
CC
    encodes a NOVX polypeptide of the invention.
XX
SO
    Sequence 2881 BP; 527 A; 985 C; 867 G; 502 T; 0 U; 0 Other;
 Query Match
                        97.2%;
                                Score 2674.8; DB 12; Length 2881;
 Best Local Similarity
                        98.9%;
                                Pred. No. 0;
 Matches 2727; Conservative
                               0; Mismatches
                                               22; Indels
                                                                 Gaps
                                                                         3;
                                                             9;
           1 CCGCGGGGCCCCGCCCGCCCGCCCGCCCGCCCGCGCCATGGCCGTCCGGCCC 60
Qу
             42 CCGCGGGGCCCGGCCCGCCCGCCCGCCGCCGGCCATGGCCGTCCGGCCC 101
Db
QУ
          61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120
```

|   | Db | 102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161  |     |
|---|----|---|-----|
|   | Qу | 121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180  | t   |
|   | Db | 162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221  |     |
|   | Qу | 181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240  | ı   |
|   | Db | 222 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 281  |     |
|   | Qу | 241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300  | ! - |
|   | Db | 282 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 341  |     |
|   | Qу | 301 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC 360  |     |
|   | Db | 342 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGC 401  |     |
|   | Qу | 361 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 420  |     |
|   | Db | 402 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 461  |     |
|   | Qу | 421 TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 480  |     |
| • | Db | 462 TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 521  |     |
|   | Qу | 481 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 540  |     |
|   | Db | 522 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 581  |     |
|   | Qу | 541 ATCGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 600     |     |
|   | Db | 582 ATCGTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 641     |     |
|   | Qу | AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 660      |     |
|   | Db | 642 AACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 701      |     |
|   | Qy | 661 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 720  |     |
|   | Db | 702 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 761  |     |
|   | Qy | 721 ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG          |     |
|   | Db | 762 ATCGTGGCACGTCGCCGCGCGCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG            |     |
|   | Qy | 781 TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG 840  |     |
|   | Db | 822 TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCGGCTGGCAGAAACGG 881   |     |
|   | Qy | 841 AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT 900  |     |
|   | Db | 882 AGCCGGAGCTGCACCCACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT 941  |     |
|   | Qy | 901 GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG 957     |     |
|   | Db | 942 GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG 100: | 1   |
|   |    | ·   |     |

| Qу   |      | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC |      |
|------|------|--|------|
| Db   | 1002 | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC | 1061 |
| Qy   | 1018 | CCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT   | 1077 |
| Db   | 1062 | CCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT   | 1121 |
| Qy   | 1078 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC | 1137 |
| Db   | 1122 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC | 1181 |
| Qy   | 1138 | CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTCATCCTCGTTTATTGC       | 1197 |
| Db   | 1182 |  | 1241 |
| Qy   | 1198 | CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC | 1257 |
| Db   | 1242 |  | 1301 |
| Qу   | 1258 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG | 1317 |
| Db   | 1302 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG | 1361 |
| Qy . | 1318 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC    | 1377 |
| Db · | 1362 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC    | 1418 |
| Qy   | 1378 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC | 1437 |
| Db   | 1419 |  | 1478 |
| Qу   | 1438 | ACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC | 1497 |
| Db   | 1479 |  | 1538 |
| Qу   | 1498 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC | 1557 |
| Db   | 1539 |  | 1598 |
| Qу   | 1558 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT  | 1617 |
| Db   | 1599 |  | 1658 |
| Qу   | 1618 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG | 1677 |
| Db   | 1659 |  | 1718 |
| Qу   | 1678 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT | 1737 |
| Db   | 1719 |  | 1778 |
| Qy   | 1738 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1797 |
| Db   | 1779 |  | 1838 |

| Qу   | 1798 | GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTG    | 1854 |
|------|------|--|------|
| Db   | 1839 |  | 1898 |
| QУ   | 1855 | CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC | 1914 |
| Db   | 1899 | CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC | 1958 |
| QУ   | 1915 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG | 1974 |
| Db   | 1959 |  | 2018 |
| Qу   | 1975 | GCTGCCGCCAAGCCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC     | 2034 |
| Db   | 2019 | GCTGCCACCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC    | 2078 |
| Qу   | 2035 | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG | 2094 |
| Db   | 2079 | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG | 2138 |
| QY . | 2095 | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC | 2154 |
| Db   | 2139 | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC | 2198 |
| Qу   | 2155 | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG | 2214 |
| Db   | 2199 | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG | 2258 |
| Qy   | 2215 | CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC | 2274 |
| Db   | 2259 | CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC | 2318 |
| Qу   | 2275 | TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG | 2334 |
| Db   | 2319 | TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG | 2378 |
| Qу   | 2335 | CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC | 2394 |
| Db   | 2379 | CTGTGGGTGTGGCAGGGGGGGGCGACGGGCAGAGCTTCAGCATCAACATCACC        | 2438 |
| Qу   | 2395 | AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG | 2454 |
| Db   | 2439 | AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG | 2498 |
| Qу   | 2455 | GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG | 2514 |
| Db   | 2499 | GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG | 2558 |
| Qу   | 2515 | GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC | 2574 |
| Db   | 2559 | GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC | 2618 |
| Qу   | 2575 | AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGG  | 2634 |
| Db   | 2619 | AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGG  | 2678 |
| Ov   | 2635 | GAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTG                   | 2694 |

```
Db
       Qу
           Db
RESULT 9
ADH71641
    ADH71641 standard; DNA; 2881 BP.
XX
AC
    ADH71641;
XX
DT
    25-MAR-2004 (first entry)
XX
    Human gene of the invention NOV21g SEQ ID NO:537.
DΕ
XX
    ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW
    anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
    vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
    obesity; diabetes; infectious disease; metabolic syndrome X;
KW
KW
    dyslipidaemia.
XX
OS
    Homo sapiens:
XX
PN
    WO2003102155-A2.
XX
PD
    11-DEC-2003.
XX
    03-JUN-2003; 2003WO-US017430.
PF
XX
    03-JUN-2002; 2002US-0385120P.
PR
    04-JUN-2002; 2002US-0385784P.
PR
    05-JUN-2002; 2002US-0386041P.
PR
    05-JUN-2002; 2002US-0386047P.
PR
    06-JUN-2002; 2002US-0386376P.
PR
    06-JUN-2002; 2002US-0386453P.
PR
    06-JUN-2002; 2002US-0386864P.
PR
    06-JUN-2002; 2002US-0387016P.
PR
    07-JUN-2002; 2002US-0386796P.
PR
    07-JUN-2002; 2002US-0386816P.
PR
    07-JUN-2002; 2002US-0386931P.
PR
    07-JUN-2002; 2002US-0386942P.
PR
    07-JUN-2002; 2002US-0386971P.
PR 1
    07-JUN-2002; 2002US-0387262P.
PR
    08-JUN-2002; 2002US-0296960P.
PR
    10-JUN-2002; 2002US-0387400P.
PR
    10-JUN-2002; 2002US-0387535P.
PR
    11-JUN-2002; 2002US-0387610P.
PR
    11-JUN-2002; 2002US-0387625P.
PR
    11-JUN-2002; 2002US-0387634P.
PR
    11-JUN-2002; 2002US-0387668P.
PR
    11-JUN-2002; 2002US-0387696P.
PR
    11-JUN-2002; 2002US-0387702P.
PR
PR
    11-JUN-2002; 2002US-0387836P.
PR
    11-JUN-2002; 2002US-0387859P.
```

```
12-JUN-2002; 2002US-0387933P.
PR
     12-JUN-2002; 2002US-0387934P.
PR
PR
     12-JUN-2002; 2002US-0387960P.
PR
     12-JUN-2002; 2002US-0388022P.
PR
     12-JUN-2002; 2002US-0388096P.
PR
     13-JUN-2002; 2002US-0389123P.
PR
     14-JUN-2002; 2002US-0389118P.
     14-JUN-2002; 2002US-0389120P.
PR
PR
     14-JUN-2002; 2002US-0389144P.
     14-JUN-2002; 2002US-0389146P.
PR
     17-JUN-2002; 2002US-0389729P.
PR
PR
     17-JUN-2002; 2002US-0389742P.
PR
     18-JUN-2002; 2002US-0389884P.
     19-JUN-2002; 2002US-0390006P.
PR
     19-JUN-2002; 2002US-0390209P.
PR
     21-JUN-2002; 2002US-0390763P.
PR
     17-JUL-2002; 2002US-0396706P.
PR
PR
     06-AUG-2002; 2002US-0401628P.
PR
     09-AUG-2002; 2002US-0402156P.
PR
     09-AUG-2002; 2002US-0402256P.
     09-AUG-2002; 2002US-0402389P.
PR
     12-AUG-2002; 2002US-0402786P.
PR
     12-AUG-2002; 2002US-0402816P.
PR
     12-AUG-2002; 2002US-0402821P.
PR
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
PR
PR
     13-AUG-2002; 2002US-0403459P.
PR
     13-AUG-2002; 2002US-0403531P.
PR
     13-AUG-2002; 2002US-0403532P.
     13-AUG-2002; 2002US-0403563P.
PR
     13-AUG-2002; 2002US-0406317P.
PR
     15-AUG-2002; 2002US-0403617P.
PR
     26-AUG-2002; 2002US-0406182P.
PR
PR
     26-AUG-2002; 2002US-0406355P.
PR
     27-AUG-2002; 2002US-0406240P.
PR
     12-SEP-2002; 2002US-0410084P.
     20-SEP-2002; 2002US-0412528P.
PR
     23-SEP-2002; 2002US-0412731P.
PR
PR
     30-SEP-2002; 2002US-0414801P.
PR
     30-SEP-2002; 2002US-0414839P.
PR
     30-SEP-2002; 2002US-0414840P.
PR
     30-SEP-2002; 2002US-0414954P.
PR
     09-OCT-2002; 2002US-0417186P.
     09-OCT-2002; 2002US-0417406P.
PR
     23-OCT-2002; 2002US-0420639P.
PR
     28-OCT-2002; 2002US-0421156P.
PR
     31-OCT-2002; 2002US-0422690P.
PR
PR
     01-NOV-2002; 2002US-0423130P.
PR
     05-NOV-2002; 2002US-00423798.
     05-NOV-2002; 2002US-0423798P.
PR
PR
     12-NOV-2002; 2002US-0425453P.
XX
PΑ
     (CURA-) CURAGEN CORP.
XX
PI.
     Alsobrook JP, Alvarez E,
                                Anderson DW, Boldog FL,
                                                           Casman SJ;
PI
     Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
ΡI
     Ettenberg S,
                   Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
```

```
ΡI
    Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
    Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PΙ
PΙ
    Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
    Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PΙ
PΙ
    Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PΙ
    Zhong H;
XX
DR
    WPI; 2004-081935/08.
DR
    P-PSDB; ADH71642.
XX
PT
    New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT
    treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
    obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS
    Example 21; SEQ ID NO 537; 1880pp; English.
XX
CC
    The invention relates to a novel isolated polypeptide (NOVX). A
    polypeptide of the invention has cytostatic, immunomodulator,
CC
    neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
CC
    antilipaemic activity, and may have a use in gene therapy, and as a
CC
    vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC
    any of the 303 fully defined nucleotide sequences given in the
    specification. The polypeptide is useful in the manufacture of a
CC
    medicament for treating a syndrome associated with a human disease. The
CC
    polypeptide, polynucleotide and antibody are useful in diagnosing,
CC
    treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC
CC
    Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC
    diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC
    further used as hybridisation probes, in chromosome mapping, tissue
CC
    typing, preventive medicine, and pharmacogenomics. The present sequence
CC
    encodes a NOVX polypeptide of the invention.
XX
SQ
    Sequence 2881 BP; 526 A; 986 C; 867 G; 502 T; 0 U; 0 Other;
 Query Match
                       97.2%; Score 2674.8; DB 12; Length 2881;
 Best Local Similarity
                       98.9%; Pred. No. 0;
 Matches 2727; Conservative
                             0; Mismatches
                                                                     3;
                                             22; Indels
                                                           9; Gaps
          1 CCGCGGGGCCCGGCCCGCCCGCCCGCCGCCGCGCGCCATGGCCGTCCGGCCC 60
Qу
            42 CCGCGGGCCCCGCCCCGCCCGCCCGCCCGCCCGCCGCCATGGCCGTCCGGCCC 101
Db
         61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120
Qy
            102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161
Db
Qу
        121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180
            162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221
Db
        181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240
Qу
            222 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 281
Db
        241 AAGGCCGTGCCCGCCAGGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300
Qу
```

282 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGGAGTGGGTGCGCCAGGTG 341

Db

|   | Qу   | 301  | GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC         | 360  |
|---|------|------|--|------|
|   | Db   | 342  | GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGC         | 401  |
|   | Qу   | 361  | ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG         | 420  |
|   | Db   | 402  | ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG         | 461  |
|   | Qy   | 421  | TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC         | 480  |
|   | Db   | 462  | TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC         | 521  |
|   | Qy   | 481  | AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC         | 540  |
|   | Db   | 522  | AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC         | 581  |
| • | Qy   | 541  | ATCGTGCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG              | 600  |
|   | Db · | 582  | ATCGTGCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG              | 641  |
|   | Qу   | 601  | AACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC             | 660  |
|   | Db   | 642  | AACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC             | 701  |
|   | Qy   | 661  | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC         | 720  |
|   | Db   | 702  | $\tt CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC$   | 761  |
|   | Qy   | 721  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG             | 780  |
|   | Db   | 762  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG             | 821  |
|   | Qy   | 781  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG         | 840  |
|   | Db   | 822  | ${\tt TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG}$ | 881  |
| • | Qу   | 841  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT         | 900  |
|   | Db   | 882  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT         | 941  |
|   | Qу   |      | GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG            |      |
|   | Db   | 942  | GTCCATGACCGCACCGTCTCTCTCTGTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG           | 1001 |
|   | Qу   | 958  | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC         | 1017 |
| • | Db   | 1002 | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC         | 1061 |
|   | Qy   | 1018 | CCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT           | 1077 |
|   | Db   | 1062 | CCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT           | 1121 |
|   | Qy   | 1078 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC         | 1137 |
|   | Db   | 1122 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC         | 1181 |

|   | Qу   | 1138 | CTCATCGCCGTGGCCGTCTGCTGGTCCTGCTGCTGCTCCTCATCCTCGTTTATTGC          | 1197 |
|---|------|------|---|------|
|   | Db   | 1182 | CTCATCGCCGTGGCCGTCTGCCTGCTGCTGCTGCTCATCCTCGTTTATTGC               | 1241 |
|   | Qу   |      | CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC      | 1257 |
|   | Db   |      | CGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC       | 1301 |
|   | Qу   | 1258 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG      | 1317 |
|   | Db   | 1302 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG      | 1361 |
|   | Qy   | 1318 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC         | 1377 |
| • | .Db  | 1362 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC         | 1418 |
|   | Qy   | 1378 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC      | 1437 |
|   | Db   | 1419 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC      | 1478 |
|   | Qy   | 1438 | ACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC      | 1497 |
|   | Db   | 1479 | ACACTGCACCACACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC               | 1538 |
|   | Qу   | 1498 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC      | 1557 |
|   | Db   | 1539 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC      | 1598 |
|   | Qy   | 1558 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT       | 1617 |
| • | Db   | 1599 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT       | 1658 |
|   | Qу   | 1618 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG      | 1677 |
|   | Db   | 1659 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG      | 1718 |
|   | Qу   | 1678 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT      | 1737 |
|   | Db   | 1719 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT      | 1778 |
|   | Qy   | 1738 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT      | 1797 |
|   | Db . | 1779 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT      | 1838 |
|   | Qy   | 1798 | GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTG         | 1854 |
|   | Db . | 1839 |   | 1898 |
|   | Qy   | 1855 | $\tt CTGCACCTGGGCGAGGGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC$ | 1914 |
|   | Db   | 1899 |   | 1958 |
|   | QУ   | 1915 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG      | 1974 |
|   | Db   | 1959 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG      | 2018 |
|   | 017  | 1075 | CCTCCCCCAACCCCCTAAACCTCCTTCCTTCTTTTCCCCCC                         | 2034 |

.

| Db      | 2019 |   |
|---------|------|---|
| Qу      | 2035 | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG 2094 |
| Db      | 2079 | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGCTGGTGCAG 2138 |
| Qy      | 2095 | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC 2154 |
| Db      | 2139 | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC 2198 |
| Qу      | 2155 | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG 2214 |
| Db      | 2199 | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG 2258 |
| Qy      | 2215 | CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC 2274 |
| Db      | 2259 | CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC 2318 |
| Qy      | 2275 | TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG 2334 |
| Db      | 2319 | TTGCACTGCACCTTGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG 2378        |
| Qġ      | 2335 | CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC 2394 |
| Db      | 2379 | CTGTGGGTGTGGCAGGGGGGGGGGGGGGGGGGGGGGGGG                           |
| Qу      | 2395 | AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2454 |
| Db      | 2439 | AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2498 |
| Qy .    | 2455 | GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2514 |
| Db      | 2499 | GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2558 |
| Qу      | 2515 | GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2574 |
| Db      | 2559 | GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2618 |
| Qy<br>: | 2575 | AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2634 |
| Db      | 2619 | AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGG 2678  |
| Qy      | 2635 | GAGGCGCGCACTTCCCCAACGCCAACCTCAGCCAGCTGGCTG                        |
| Db      | 2679 | GAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTG                        |
| Qy      | 2695 | GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCC          |
| Db      | 2739 | GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCC          |

## RESULT 10 ADH71629

ID ADH71629 standard; DNA; 2881 BP.

XX

AC ADH71629;

```
XX
DT
     25-MAR-2004
                  (first entry)
XX
     Human gene of the invention NOV21k SEQ ID NO:525.
DE
XX
     ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW
     anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
     vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
KW
     obesity; diabetes; infectious disease; metabolic syndrome X;
KW
     dyslipidaemia.
XX
os
     Homo sapiens.
XX
PN
     WO2003102155-A2.
XX
PD
     11-DEC-2003.
XX
ΡF
     03-JUN-2003; 2003WO-US017430.
XX
PR
     03-JUN-2002; 2002US-0385120P.
PR
     04-JUN-2002; 2002US-0385784P.
     05-JUN-2002; 2002US-0386041P.
PR
     05-JUN-2002; 2002US-0386047P.
PR
     06-JUN-2002; 2002US-0386376P.
PR
     06-JUN-2002; 2002US-0386453P.
PR
     06-JUN-2002; 2002US-0386864P.
PR
PR
     06-JUN-2002; 2002US-0387016P.
PR
     07-JUN-2002; 2002US-0386796P.
PR
     07-JUN-2002; 2002US-0386816P.
     07-JUN-2002; 2002US-0386931P.
PR
     07-JUN-2002; 2002US-0386942P.
PR
     07-JUN-2002; 2002US-0386971P.
PR
PR
     07-JUN-2002; 2002US-0387262P.
     08-JUN-2002; 2002US-0296960P.
PR
PR
     10-JUN-2002; 2002US-0387400P.
     10-JUN-2002; 2002US-0387535P.
PR
     11-JUN-2002; 2002US-0387610P.
PR
     11-JUN-2002; 2002US-0387625P.
PR
     11-JUN-2002; 2002US-0387634P.
PR
PR
     11-JUN-2002; 2002US-0387668P.
     11-JUN-2002; 2002US-0387696P.
PR
PR
     11-JUN-2002; 2002US-0387702P.
     11-JUN-2002; 2002US-0387836P.
PR
     11-JUN-2002; 2002US-0387859P.
PR
     12-JUN-2002; 2002US-0387933P.
PR
     12-JUN-2002; 2002US-0387934P.
PR
PR
     12-JUN-2002; 2002US-0387960P.
PR
     12-JUN-2002; 2002US-0388022P.
PR
     12-JUN-2002; 2002US-0388096P.
     13-JUN-2002; 2002US-0389123P.
PR
     14-JUN-2002; 2002US-0389118P.
PR
     14-JUN-2002; 2002US-0389120P.
PR
PR
     14-JUN-2002; 2002US-0389144P.
PR
     14-JUN-2002; 2002US-0389146P.
PR
     17-JUN-2002; 2002US-0389729P.
PR
     17-JUN-2002; 2002US-0389742P.
```

PR

18-JUN-2002; 2002US-0389884P.

```
PR
     19-JUN-2002; 2002US-0390006P.
     19-JUN-2002; 2002US-0390209P.
PR
     21-JUN-2002; 2002US-0390763P.
PR
     17-JUL-2002; 2002US-0396706P.
PR
     06-AUG-2002; 2002US-0401628P.
PR
     09-AUG-2002; 2002US-0402156P.
PR
     09-AUG-2002; 2002US-0402256P.
PR
     09-AUG-2002; 2002US-0402389P.
PR
     12-AUG-2002; 2002US-0402786P.
PR
PR
     12-AUG-2002; 2002US-0402816P.
     12-AUG-2002; 2002US-0402821P.
PR
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
PR
     13-AUG-2002; 2002US-0403459P.
PR
     13-AUG-2002; 2002US-0403531P.
PR
     13-AUG-2002; 2002US-0403532P.
PR
     13-AUG-2002; 2002US-0403563P.
PR
PR
     13-AUG-2002; 2002US-0406317P.
PR
     15-AUG-2002; 2002US-0403617P.
     26-AUG-2002; 2002US-0406182P.
PR
     26-AUG-2002; 2002US-0406355P.
PR
     27-AUG-2002; 2002US-0406240P.
PR
     12-SEP-2002; 2002US-0410084P.
PR
     20-SEP-2002; 2002US-0412528P.
PR
     23-SEP-2002; 2002US-0412731P.
PR
PR
     30-SEP-2002; 2002US-0414801P.
PR
     30-SEP-2002; 2002US-0414839P.
     30-SEP-2002; 2002US-0414840P.
PR
     30-SEP-2002; 2002US-0414954P.
PR
     09-OCT-2002; 2002US-0417186P.
PR
PR
     09-OCT-2002; 2002US-0417406P.
PR
     23-OCT-2002; 2002US-0420639P.
PR
     28-OCT-2002; 2002US-0421156P.
PR
     31-OCT-2002; 2002US-0422690P.
     01-NOV-2002; 2002US-0423130P.
PR
     05-NOV-2002; 2002US-00423798.
PR
     05-NOV-2002; 2002US-0423798P.
PR
     12-NOV-2002; 2002US-0425453P.
PR
XX
PA
     (CURA-) CURAGEN CORP.
XX
PΙ
     Alsobrook JP, Alvarez E, Anderson DW, Boldog FL,
                                                          Casman SJ;
PΙ
     Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
     Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PI
     Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI
PΙ
     Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PΙ
     Padigaru M,
                  Patturajan M,
                                 Pena CEA, Peyman JA, Raha D,
                                                                 Rastelli L;
ΡI
     Rieger DK, Rothenberg ME,
                                 Sciore P, Shenoy SG, Shimkets RA;
PΙ
     Smithson G,
                  Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PΙ
     Zhong H;
XX
     WPI; 2004-081935/08.
DR
DR
     P-PSDB; ADH71630.
XX
PT
    New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT
     treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
     obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
```

```
XX
PS
    Example 21; SEQ ID NO 525; 1880pp; English.
XX
CC
    The invention relates to a novel isolated polypeptide (NOVX). A
    polypeptide of the invention has cytostatic, immunomodulator,
CC
CC
    neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
    antilipaemic activity, and may have a use in gene therapy, and as a
CC
    vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
    any of the 303 fully defined nucleotide sequences given in the
CC
CC
    specification. The polypeptide is useful in the manufacture of a
    medicament for treating a syndrome associated with a human disease. The
CC
    polypeptide, polynucleotide and antibody are useful in diagnosing,
CC
    treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC
CC
    Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
    diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC
    further used as hybridisation probes, in chromosome mapping, tissue
CC
    typing, preventive medicine, and pharmacogenomics. The present sequence
CC
    encodes a NOVX polypeptide of the invention.
CC
XX
    Sequence 2881 BP; 526 A; 986 C; 868 G; 501 T; 0 U; 0 Other;
SO
                       97.28;
                              Score 2674.8; DB 12; Length 2881;
 Query Match
                       98.9%;
 Best Local Similarity
                              Pred. No. 0;
 Matches 2727: Conservative
                             0: Mismatches
                                            22:
                                                Indels
                                                          9:
                                                             Gaps
                                                                    3;
          1 CCGCGGGCCCCGCCCGCCCGCCCGCCGCCGCGCCATGGCCGTCCGGCCC 60
Qу
            Db
          42 CCGCGGGGCCCGCCCGCCCGCCCGCCGCCGCGGCCATGGCCGTCCGGCCC 101
          61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120
Qy
            102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161
Db
         121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180
Qу
            162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221
Db
         181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240
Qy
            222 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 281
Db
         241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300
Qy
```

301 GACCACGTGATCGAGCGCAGCAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC 360

342 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGC 401

361 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 420

421 TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC 480

Db

Qу

Db

Qу

Db

Qу

Db

| Qy       | 481 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 540  |
|----------|---|
| Db       | 522 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 581  |
| Qy       | 541 ATCGTGCTGCCCTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 600   |
| Db       | 582 ATCGTGCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 641   |
| Qy       | 601 AACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 660  |
| Db       | 642 AACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 701  |
| Qy       | 661 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 720  |
| Db       | 702 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 761  |
| Qy       | 721 ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTACGTGAACGGTGGGTG   |
| Db       | 762 ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG  |
| Qy       | 781 TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG 840  |
| Db       | 822 TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG 881  |
| Qy       | 841 AGCCGGAGCTGCACCCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT 900   |
| Db       | 882 AGCCGGAGCTGCACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT 941  |
| Qу       | 901 GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG 957   |
| Db       | 942 GTCCATGACCGCACCGTCTCTCTCTGTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG 1001   |
| Qy       | 958 AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC 1017   |
| Db       | 1002 AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC 1061  |
| Qy       | 1018 CCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT 1077   |
| Db       | 1062 CCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT 1121  1078 ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC 1137 |
| Qy<br>Db |   |
| Qy       | 1138 CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTCTCTCTCTCTCTTTATTGC 1197   |
| ₽₽       |   |
| Qy       | 1198 CGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC 1257   |
| . Db     |   |
| Qy       | 1258 CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACACCCCCATCTGCTCACCATCCAGCCG 1317   |
| Db       |   |
| Ov       | 1318 GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC 1377   |

.

|    |      | ***************************************                      |      |
|----|------|--|------|
| Db | 1362 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC    | 1418 |
| Qу | 1378 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC | 1437 |
| Db | 1419 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC | 1478 |
| Qу | 1438 | ACACTGCACCACGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC   | 1497 |
| Db | 1479 | ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC  | 1538 |
| Qy | 1498 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC | 1557 |
| Db | 1539 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC | 1598 |
| Qу | 1558 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT  | 1617 |
| Db | 1599 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT  | 1658 |
| Qу | 1618 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG | 1677 |
| Db | 1659 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG | 1718 |
| Qу | 1678 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT | 1737 |
| Db | 1719 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT | 1778 |
| Qу | 1738 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1797 |
| Db | 1779 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1838 |
| Qу | 1798 | GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTG    | 1854 |
| Db | 1839 | GACAGCTGGAGCCTCCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG      | 1898 |
| Qу | 1855 | CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC | 1914 |
| Db | 1899 | CTGCACCTGGGCGAGGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC   | 1958 |
| Qу | 1915 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG | 1974 |
| Db | 1959 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG | 2018 |
| Qy | 1975 | GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC    | 2034 |
| Db | 2019 | GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC    | 2078 |
| Qу | 2035 | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG | 2094 |
| Db | 2079 | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG | 2138 |
| Qу | 2095 | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC | 2154 |
| Db | 2139 | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC | 2198 |
| Qу | 2155 | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG | 2214 |
|    |      |  |      |

```
2199 AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG 2258
Db
     2215 CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC 2274
Qу
         2259 CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC 2318
Db
     2275 TTGCACTGCACCTTGACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG 2334
Qу
         2319 TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG 2378
Db
      2335 CTGTGGGTGTGGCAGGTGGAGGGCGACGGCAGAGCTTCAGCATCAACTTCAACATCACC 2394
Qy
         2379 CTGTGGGTGTGGCAGGTGGAGGGCGACGGCAGAGCTTCAGCATCAACTTCAACATCACC 2438
Db
      2395 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2454
Qy
         2439 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2498
Db
      2455 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2514
Qу
         2499 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2558
Db
      2515 GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2574
Qy
         2559 GACCCACCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2618
Db
      2575 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGG 2634
Qу
         2619 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGG 2678
Db
      Qу
         Db
      Qу
         Db
RESULT 11
ADH71631
   ADH71631 standard; DNA; 2881 BP.
XX
AC
   ADH71631;
XX
DT
   25-MAR-2004 (first entry)
XX
DE
   Human gene of the invention NOV211 SEQ ID NO:527.
XX
KW
   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
   vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
   obesity; diabetes; infectious disease; metabolic syndrome X;
KW
KW
   dyslipidaemia.
XX
```

os

XX

Homo sapiens.

```
WO2003102155-A2.
PN
XX
    11-DEC-2003.
PD
XX
     03-JUN-2003; 2003WO-US017430.
PF
XX
     03-JUN-2002; 2002US-0385120P.
PR
     04-JUN-2002; 2002US-0385784P.
PR
     05-JUN-2002; 2002US-0386041P.
PR
     05-JUN-2002; 2002US-0386047P.
PR
     06-JUN-2002; 2002US-0386376P.
PR
PR
     06-JUN-2002; 2002US-0386453P.
     06-JUN-2002; 2002US-0386864P.
PR
     06-JUN-2002; 2002US-0387016P.
PR
     07-JUN-2002; 2002US-0386796P.
PR
     07-JUN-2002; 2002US-0386816P.
PR
PR
     07-JUN-2002; 2002US-0386931P.
     07-JUN-2002; 2002US-0386942P.
PR
PR
     07-JUN-2002; 2002US-0386971P.
     07-JUN-2002; 2002US-0387262P.
PR
     08-JUN-2002; 2002US-0296960P.
PR
     10-JUN-2002; 2002US-0387400P.
PR
     10-JUN-2002; 2002US-0387535P.
PR
     11-JUN-2002; 2002US-0387610P.
PR
     11-JUN-2002; 2002US-0387625P.
PR
     11-JUN-2002; 2002US-0387634P.
PR
PR
     11-JUN-2002; 2002US-0387668P.
PR
     11-JUN-2002; 2002US-0387696P.
     11-JUN-2002; 2002US-0387702P.
PR
     11-JUN-2002; 2002US-0387836P.
PR
     11-JUN-2002; 2002US-0387859P.
PR
     12-JUN-2002; 2002US-0387933P.
PR
PR
     12-JUN-2002; 2002US-0387934P.
     12-JUN-2002; 2002US-0387960P.
PR
PR
     12-JUN-2002; 2002US-0388022P.
PR
     12-JUN-2002; 2002US-0388096P.
     13-JUN-2002; 2002US-0389123P.
PR
     14-JUN-2002; 2002US-0389118P.
PR
PR
     14-JUN-2002; 2002US-0389120P.
PR
     14-JUN-2002; 2002US-0389144P.
PR
     14-JUN-2002; 2002US-0389146P.
     17-JUN-2002; 2002US-0389729P.
PR
     17-JUN-2002; 2002US-0389742P.
PR
     18-JUN-2002; 2002US-0389884P.
PR
     19-JUN-2002; 2002US-0390006P.
PR
     19-JUN-2002; 2002US-0390209P.
PR
     21-JUN-2002; 2002US-0390763P.
PR
PR
     17-JUL-2002; 2002US-0396706P.
     06-AUG-2002; 2002US-0401628P.
PR
     09-AUG-2002; 2002US-0402156P.
PR
     09-AUG-2002; 2002US-0402256P.
PR
     09-AUG-2002; 2002US-0402389P.
PR
PR
     12-AUG-2002; 2002US-0402786P.
     12-AUG-2002; 2002US-0402816P.
PR
PR
     12-AUG-2002; 2002US-0402821P.
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
PR
```

```
13-AUG-2002; 2002US-0403459P.
PR
    13-AUG-2002; 2002US-0403531P.
PR
    13-AUG-2002; 2002US-0403532P.
PR
    13-AUG-2002; 2002US-0403563P.
PR
PR
    13-AUG-2002; 2002US-0406317P.
    15-AUG-2002; 2002US-0403617P.
PR
    26-AUG-2002; 2002US-0406182P.
PR
    26-AUG-2002; 2002US-0406355P.
PR
    27-AUG-2002; 2002US-0406240P.
PR
    12-SEP-2002; 2002US-0410084P.
PR
PR
    20-SEP-2002; 2002US-0412528P.
    23-SEP-2002; 2002US-0412731P.
PR
    30-SEP-2002; 2002US-0414801P.
PR
     30-SEP-2002; 2002US-0414839P.
PR
     30-SEP-2002; 2002US-0414840P.
PR
    30-SEP-2002; 2002US-0414954P.
PR
    09-OCT-2002; 2002US-0417186P.
PR
PR
    09-OCT-2002; 2002US-0417406P.
    23-OCT-2002; 2002US-0420639P.
PR
    28-OCT-2002; 2002US-0421156P.
PR
     31-OCT-2002; 2002US-0422690P.
PR
PR
     01-NOV-2002; 2002US-0423130P.
     05-NOV-2002; 2002US-00423798.
PR
     05-NOV-2002; 2002US-0423798P.
PR
PR
     12-NOV-2002; 2002US-0425453P.
XX
PΑ
     (CURA-) CURAGEN CORP.
XX
    Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
ΡI
     Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PΙ
     Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
ΡI
     Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PΙ
    Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PΙ
     Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;
PΙ
     Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PΙ
     Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
ΡI
PI
     Zhong H;
XX
DR
    WPI; 2004-081935/08.
DR
     P-PSDB; ADH71632.
XX
    New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT
PT
     treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
     obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
PT
XX
PS
     Example 21; SEQ ID NO 527; 1880pp; English.
XX
CC
     The invention relates to a novel isolated polypeptide (NOVX). A
CC
     polypeptide of the invention has cytostatic, immunomodulator,
CC
     neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
     antilipaemic activity, and may have a use in gene therapy, and as a
CC
     vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC
     any of the 303 fully defined nucleotide sequences given in the
CC
CC
     specification. The polypeptide is useful in the manufacture of a
CC
     medicament for treating a syndrome associated with a human disease. The
CC
     polypeptide, polynucleotide and antibody are useful in diagnosing,
CC
     treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
```

```
diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC
CC
   further used as hybridisation probes, in chromosome mapping, tissue
   typing, preventive medicine, and pharmacogenomics. The present sequence
CC
   encodes a NOVX polypeptide of the invention.
CC
XX
   Sequence 2881 BP; 526 A; 984 C; 868 G; 503 T; 0 U; 0 Other;
SQ
                   97.2%;
                         Score 2674.8;
                                    DB 12;
                                           Length 2881;
 Query Match
                   98.9%;
 Best Local Similarity
                         Pred. No. 0;
                                         Indels
                                                 9;
 Matches 2727; Conservative
                           Mismatches
                                     22:
                                                          3;
                        0;
                                                    Gaps
         1 CCGCGGGGCCCGGCCCGCCCGCCGCCGCCGCGCGCATGGCCGTCCGGCCC 60
Qy
          42 CCGCGGGGCCCGGCCCGGCCCGCCTGCCCGCGCGGCCATGGCCGTCCGGCCC 101
Db
        61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120
Qγ
          102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161
Db
       121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180
Qу
          162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221
Db
       181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240
Qу
          222 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 281
Db
       241 AAGGCCGTGCCCGCCAGGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300
Qy
          282 AAGGCCGTGCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 341
Db
       301 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC 360
Qу
          342 GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGC 401
Db
       361 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 420
Qy
          402 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 461
Db
       421 TGCGTGGCATGGAGCTCCTCGGGCACCACAAGAGTCAGAAGGCCTACATCCGCATAGCC 480
Qy
          462 TGCGTGGCATGGAGCTCCTCGGGCACCACAAGAGTCAGAAGGCCTACATCCGCATAGCC 521
Db
       481 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 540
Qу
          522 AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC 581
Db
       541 ATCGTGCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 600
Qy
          582 ATCGTGCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG 641
Db
       601 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 660
Qу
          642 AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC 701
Db
```

661 CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC 720

Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious

CC

Qу

| Db | 702  |  | 761   |
|----|------|--|-------|
| Qу | 721  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG     | 780   |
| Db | 762  |  | 821   |
| Qy | 781  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG | 840   |
| Db | 822  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCCGGCTGGCAGAAACGG  | 881   |
| Qy | 841  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT | 900   |
| Db | 882  |  | 941   |
| Qy | 901  | GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG    | 957   |
| Db | 942  | GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG | 1001  |
| Qу | 958  | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC | 1017  |
| Db | 1002 | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC | 1061  |
| Qy | 1018 | CCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT   | 1077  |
| Db | 1062 | CCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT  | 1121  |
| Qу | 1078 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC | 1137  |
| Db | 1122 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC | 1181  |
| Qу | 1138 | CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTCATCCTCGTTTATTGC       | 1197  |
| Db | 1182 | CTCATCGCCGTGGCCTGCTGCTGCTGCTGCTGCTCATCCTCGTTTATTGC           | 1241  |
| QУ | 1198 | CGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC  | 1257  |
| Db | 1242 | CGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC  | 1301  |
| Qу | 1258 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG | 1317  |
| Db | 1302 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACACCCCCATCTGCTCACCATCCAGCCG  | 1361  |
| QУ | 1318 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC    | 1377  |
| Db | 1362 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC    | 1418  |
| Qy | 1378 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC | 1437  |
| Db | 1419 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC | 1478  |
| Qу | 1438 | ACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC | 149.7 |
| Ďр | 1479 | ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC  | 1538  |
| Qу | 1498 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC | 1557  |

| Db | 1539  | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC | 1598 |
|----|-------|--|------|
| Qу | 1558  | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT  | 1617 |
| Db | 1599  | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT  | 1658 |
| Qу | 1618  | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG | 1677 |
| Db | 1659  | GTCATACCCCGAGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG  | 1718 |
| Qу | 1678  | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT | 1737 |
| Db | 1719  | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT | 1778 |
| Qу | 173,8 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1797 |
| Db | 1779  | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1838 |
| Qу | 1798  | GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTG    | 1854 |
| Db | 1839  | GACAGCTGGAGCCTCCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG      | 1898 |
| Qу | 1855  | CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC | 1914 |
| Db | 1899  | CTGCACCTGGGCGAGGAGGCCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC  | 1958 |
| QУ | 1915  | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG | 1974 |
| Db | 1959  | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG | 2018 |
| Qу | 1975  | GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC    | 2034 |
| Db | 2019  | GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC    | 2078 |
| Qу | 2035  | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG | 2094 |
| Db | 2079  | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG | 2138 |
| Qу | 2095  | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC | 2154 |
| Db | 2139  | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC | 2198 |
| Qу | 2155  | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG | 2214 |
| Db | 2199  | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG | 2258 |
| Qу | 2215  | CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC | 2274 |
| Db | 2259  | CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC | 2318 |
| Qу | 2275  | TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG | 2334 |
| Db | 2319  | TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG | 2378 |
| Qу | 2335  | CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC | 2394 |
| Db | 2379  | CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC | 2438 |

```
2395 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2454
Qу
         2439 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2498
Db
      2455 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2514
Qу
          2499 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2558
Db
      2515 GACCCACCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2574
Qу
          2559 GACCCACCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2618
Db
      2575 AGCCATCTCAGCTTCTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGG 2634
Qy
          2619 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2678
Db
      Qу
          Db
      Qу
          Db
RESULT 12
ADH71645
   ADH71645 standard; DNA; 2881 BP.
XX
AC
   ADH71645;
XX
DT
   25-MAR-2004
            (first entry)
XX
DE
   Human gene of the invention NOV21s SEQ ID NO:541.
XX
   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW
   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
   vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
   obesity; diabetes; infectious disease; metabolic syndrome X;
KW
KW
   dyslipidaemia.
XX
OS
   Homo sapiens.
XX
PN
   WO2003102155-A2.
XX
PD
   11-DEC-2003.
XX
   03-JUN-2003; 2003WO-US017430.
PF
XX
   03-JUN-2002; 2002US-0385120P.
PR
PR
   04-JUN-2002; 2002US-0385784P.
PR
   05-JUN-2002; 2002US-0386041P.
PR
   05-JUN-2002; 2002US-0386047P.
PR
   06-JUN-2002; 2002US-0386376P.
PR
   06-JUN-2002; 2002US-0386453P.
PR
   06-JUN-2002; 2002US-0386864P.
```

```
06-JUN-2002; 2002US-0387016P.
PR
     07-JUN-2002; 2002US-0386796P.
PR
     07-JUN-2002; 2002US-0386816P.
PR
     07-JUN-2002; 2002US-0386931P.
PR
PR
     07-JUN-2002; 2002US-0386942P.
     07-JUN-2002; 2002US-0386971P.
PR
     07-JUN-2002; 2002US-0387262P.
PR
     08-JUN-2002; 2002US-0296960P.
PR
     10-JUN-2002; 2002US-0387400P.
PR
     10-JUN-2002; 2002US-0387535P.
PR
     11-JUN-2002; 2002US-0387610P.
PR
     11-JUN-2002; 2002US-0387625P.
PR
     11-JUN-2002; 2002US-0387634P.
PR
     11-JUN-2002; 2002US-0387668P.
PR
     11-JUN-2002; 2002US-0387696P.
PR
PR
     11-JUN-2002; 2002US-0387702P.
     11-JUN-2002; 2002US-0387836P.
PR
     11-JUN-2002; 2002US-0387859P.
PR
     12-JUN-2002; 2002US-0387933P.
PR
     12-JUN-2002; 2002US-0387934P.
PR
     12-JUN-2002; 2002US-0387960P.
PR
     12-JUN-2002; 2002US-0388022P.
PR
     12-JUN-2002; 2002US-0388096P.
PR
PR
     13-JUN-2002; 2002US-0389123P.
PR
     14-JUN-2002; 2002US-0389118P.
PR
     14-JUN-2002; 2002US-0389120P.
     14-JUN-2002; 2002US-0389144P.
PR
     14-JUN-2002; 2002US-0389146P.
PR
     17-JUN-2002; 2002US-0389729P.
PR
     17-JUN-2002; 2002US-0389742P.
PR
     18-JUN-2002; 2002US-0389884P.
PR
     19-JUN-2002; 2002US-0390006P.
PR
     19-JUN-2002; 2002US-0390209P.
PR
PR
     21-JUN-2002; 2002US-0390763P.
     17-JUL-2002; 2002US-0396706P.
PR
     06-AUG-2002; 2002US-0401628P.
PR
     09-AUG-2002; 2002US-0402156P.
PR
     09-AUG-2002; 2002US-0402256P.
PR
     09-AUG-2002; 2002US-0402389P.
PR
     12-AUG-2002; 2002US-0402786P.
PR
     12-AUG-2002; 2002US-0402816P.
PR
PR
     12-AUG-2002; 2002US-0402821P.
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
PR
     13-AUG-2002; 2002US-0403459P.
PR
     13-AUG-2002; 2002US-0403531P.
PR
     13-AUG-2002; 2002US-0403532P.
PR
PR
     13-AUG-2002; 2002US-0403563P.
PR
     13-AUG-2002; 2002US-0406317P.
     15-AUG-2002; 2002US-0403617P.
PR
     26-AUG-2002; 2002US-0406182P.
PR
     26-AUG-2002; 2002US-0406355P.
PR
     27-AUG-2002; 2002US-0406240P.
PR
     12-SEP-2002; 2002US-0410084P.
PR
     20-SEP-2002; 2002US-0412528P.
PR
     23-SEP-2002; 2002US-0412731P.
PR
     30-SEP-2002; 2002US-0414801P.
PR
```

```
30-SEP-2002; 2002US-0414839P.
PR
     30-SEP-2002; 2002US-0414840P.
PR
     30-SEP-2002; 2002US-0414954P.
PR
     09-OCT-2002; 2002US-0417186P.
PR
PR
     09-OCT-2002; 2002US-0417406P.
     23-OCT-2002; 2002US-0420639P.
PR
     28-OCT-2002; 2002US-0421156P.
PR
     31-OCT-2002; 2002US-0422690P.
PR
     01-NOV-2002; 2002US-0423130P.
PR
     05-NOV-2002; 2002US-00423798.
PR
PR
     05-NOV-2002; 2002US-0423798P.
PR
     12-NOV-2002; 2002US-0425453P.
XX
PA
     (CURA-) CURAGEN CORP.
XX
PΙ
     Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
ΡI
     Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PI
     Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PΙ
     Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PΙ
     Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
     Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L; Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PΙ
PI
PΙ
     Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PI
     Zhong H;
XX
DR
     WPI; 2004-081935/08.
DR
     P-PSDB; ADH71646.
XX
PT
     New NOVX polypeptides and nucleic acid molecules useful for preventing or
     treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
PT
     obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS
     Example 21; SEQ ID NO 541; 1880pp; English.
XX
CC
     The invention relates to a novel isolated polypeptide (NOVX). A
CC
     polypeptide of the invention has cytostatic, immunomodulator,
CC
     neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
     antilipaemic activity, and may have a use in gene therapy, and as a
CC
     vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC
     any of the 303 fully defined nucleotide sequences given in the
CC
     specification. The polypeptide is useful in the manufacture of a
CC
     medicament for treating a syndrome associated with a human disease. The
CC
     polypeptide, polynucleotide and antibody are useful in diagnosing,
CC
     treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC
     Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC
     diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
     further used as hybridisation probes, in chromosome mapping, tissue
CC
CC
     typing, preventive medicine, and pharmacogenomics. The present sequence
CC
     encodes a NOVX polypeptide of the invention.
XX
SQ
     Sequence 2881 BP; 526 A; 986 C; 868 G; 501 T; 0 U; 0 Other;
  Query Match
                          97.2%;
                                   Score 2674.8; DB 12; Length 2881;
  Best Local Similarity
                          98.9%;
                                   Pred. No. 0;
 Matches 2727; Conservative
                                  0;
                                     Mismatches
                                                   22;
                                                        Indels
                                                                   9;
                                                                       Gaps
                                                                               3;
```

Qy

| Db   | 42  | CCGCGGGCCCGCCCGCCCGCCCGCCCGCCCGCCCATGGCCGTCCGGCCC             | 101 |  |
|------|-----|---|-----|--|
| Qу   | 61  | GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC  | 120 |  |
| Db   | 102 | GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC  | 161 |  |
| Qу   | 121 | CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC  | 180 |  |
| Db   | 162 | CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC  |     |  |
| Qy , | 181 | TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC  | 240 |  |
| Db   | 222 | TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC  | 281 |  |
| Qу   | 241 | AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG  | 300 |  |
| Db   | 282 | AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG  | 341 |  |
| Qу   | 301 | GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC  | 360 |  |
| Db   | 342 | GACCACGTGATCGAGCGCAGCAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGC    | 401 |  |
| QУ   | 361 | ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG  | 420 |  |
| Db   | 402 | ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG  | 461 |  |
| Qу   | 421 | TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC  | 480 |  |
| Db   | 462 | TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC  | 521 |  |
| Qу   | 481 | AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC  | 540 |  |
| Db   | 522 | AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC  | 581 |  |
| Qу   | 541 | ATCGTGCTGCCCTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG       | 600 |  |
| Db   | 582 | ATCGTGCTGCCCTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG       | 641 |  |
| Qy   | 601 | AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC  | 660 |  |
| Db   | 642 | AACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC      | 701 |  |
| Qу   | 661 | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC  | 720 |  |
| Db   | 702 | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC  | 761 |  |
| Qy   | 721 | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG      |     |  |
| Db   |     | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG      |     |  |
| Qу   | 781 | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGGCGGCTGGCAGAAACGG | 840 |  |
| Db   | 822 | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG  | 881 |  |
| Qу   | 841 | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT  | 900 |  |

| Db   | 882   | AGCCGGAGCTGCACCCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT | 941          |
|------|-------|---|--------------|
| Qу   | 901   | GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG     | 957          |
| Db   | 942   | GTCCATGACCGCACCGTCTCCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG    | 1001         |
| Qу   | 958   | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC  | 1017         |
| Db   | 1002  | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC  | 1061         |
| QΥ . | 1018  | CCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT   | 107 <b>7</b> |
| Db   | 1062  | CCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTC    | 1121         |
| Qy   | 1078  | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC  | 1137         |
| Db   | 1122  | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC  | 1181         |
| Qу   | 1138  | CTCATCGCCGTGGCCGTCTGCTGGTCCTGCTGCTGCTGTTGTCCTCATCCTCGTTTATTGC | 1197         |
| Db   | 1182  | CTCATCGCCGTGGCCGTCTGCCTGCTGCTGCTGCTTGTCCTCATCCTCGTTTATTGC     | 1241         |
| QУ   | 1198  | CGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC  | 1257         |
| Db   | 1242  | CGGAAGAAGGAGGGCTGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC    | 1301         |
| QУ   | 1258  | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG  | 1317         |
| Db   | 1302  | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG  | 1361         |
| Qу   | 1318  | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC     | 1377         |
| Db   | 1362  | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATG          | 1418         |
| QУ   | 1378  | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC  | 1437         |
| Db   | 1419  | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC  | 1478         |
| Qу   | 1.438 | ACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC  | 1497         |
| Db   | 1479  | ACACTGCACCACACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC           | 1538         |
| Qy   | 1498  | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC  | 1557         |
| Db   | 1539  | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC  | 1598         |
| Qу   | 1558  | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT   |              |
| Db   | 1599  | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT   | 1658         |
| Qу   | 1618  | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG  | 1677         |
| Db   | 1659  | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG  | 1718         |
| QУ   | 1678  | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT  | 1737         |
| Db   | 1719  | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT  | 1778         |

|   | Qу   | 1738 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1797 |
|---|------|------|--|------|
|   | Db   | 1779 |  | 1838 |
|   | Qу   | 1798 | GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGA'       | 1854 |
|   | Db   | 1839 |  | 1898 |
|   | Qy   | 1855 | CTGCACCTGGGCGAGGAGGCCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC  | 1914 |
|   | Db   | 1899 |  | 1958 |
| • | Qy   | 1915 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG | 1974 |
|   | Db   | 1959 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGT  | 2018 |
|   | Qy   | 1975 | GCTGCCGCCAAGCCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC     | 2034 |
|   | Db   | 2019 | GCTGCCGCCAAGCCTCCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC     | 2078 |
|   | Qy   | 2035 | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG | 2094 |
|   | Db   | 2079 | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG | 2138 |
|   | Qy   | 2095 | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC | 2154 |
|   | Db   | 2139 | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC | 2198 |
|   | Qу   | 2155 | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG | 2214 |
| • | Db   | 2199 | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG | 2258 |
|   | Qy   | 2215 | CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC | 2274 |
|   | Db   | 2259 | CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC | 2318 |
|   | Qу   | 2275 | TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG | 2334 |
|   | Db   | 2319 | TTGCACCGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG | 2378 |
|   | Qу   | 2335 | CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC | 2394 |
|   | Db   | 2379 | CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTCACC        | 2438 |
|   | Qy · |      | AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG | 2454 |
|   | Db   |      | AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG | 2498 |
|   | Qy   | 2455 | GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG | 2514 |
|   | Db   | 2499 | GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG | 2558 |
|   | Qy   | 2515 | GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC | 2574 |
|   | Db   | 2559 |  | 2618 |

```
2575 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2634
Qу
          2619 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2678
Db
      Qу
          Db
       Qу
          Db
RESULT 13
ADH71627
   ADH71627 standard; DNA; 2881 BP.
ID
XX
AC
   ADH71627;
XX
DT
   25-MAR-2004 (first entry)
XX
DE
   Human gene of the invention NOV21j SEQ ID NO:523.
XX
KW
   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
   vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
KW
   obesity; diabetes; infectious disease; metabolic syndrome X;
KW
   dyslipidaemia.
XX
OS
   Homo sapiens.
XX
PN
   WO2003102155-A2.
XX
PD
   11-DEC-2003.
XX
PF
   03-JUN-2003; 2003WO-US017430.
XX
   03-JUN-2002; 2002US-0385120P.
PR
   04-JUN-2002; 2002US-0385784P.
PR
   05-JUN-2002; 2002US-0386041P.
PR
   05-JUN-2002; 2002US-0386047P.
PR
   06-JUN-2002; 2002US-0386376P.
PR
   06-JUN-2002; 2002US-0386453P.
PR
   06-JUN-2002; 2002US-0386864P.
PR
   06-JUN-2002; 2002US-0387016P.
PR
   07-JUN-2002; 2002US-0386796P.
PR
PR
   07-JUN-2002; 2002US-0386816P.
PR
   07-JUN-2002; 2002US-0386931P.
PR
   07-JUN-2002; 2002US-0386942P.
   07-JUN-2002; 2002US-0386971P.
PR
   07-JUN-2002; 2002US-0387262P.
PR
   08-JUN-2002; 2002US-0296960P.
PR
   10-JUN-2002; 2002US-0387400P.
PR
   10-JUN-2002; 2002US-0387535P.
PR
PR
   11-JUN-2002; 2002US-0387610P.
PR
   11-JUN-2002; 2002US-0387625P.
   11-JUN-2002; 2002US-0387634P.
PR
```

```
11-JUN-2002; 2002US-0387668P.
PR
     11-JUN-2002; 2002US-0387696P.
PR
PR
     11-JUN-2002; 2002US-0387702P.
PR
     11-JUN-2002; 2002US-0387836P.
PR
     11-JUN-2002; 2002US-0387859P.
PR
     12-JUN-2002; 2002US-0387933P.
     12-JUN-2002; 2002US-0387934P.
PR
     12-JUN-2002; 2002US-0387960P.
PR
PR
     12-JUN-2002; 2002US-0388022P.
PR
     12-JUN-2002; 2002US-0388096P.
     13-JUN-2002; 2002US-0389123P.
PR
     14-JUN-2002; 2002US-0389118P.
PR
PR
     14-JUN-2002; 2002US-0389120P.
     14-JUN-2002; 2002US-0389144P.
PR
     14-JUN-2002; 2002US-0389146P.
PR
     17-JUN-2002; 2002US-0389729P.
PR
     17-JUN-2002; 2002US-0389742P.
PR
     18-JUN-2002; 2002US-0389884P.
PR
     19-JUN-2002; 2002US-0390006P.
PR
     19-JUN-2002; 2002US-0390209P.
PR
     21-JUN-2002; 2002US-0390763P.
PR
     17-JUL-2002; 2002US-0396706P.
PR
     06-AUG-2002; 2002US-0401628P...
PR
     09-AUG-2002; 2002US-0402156P.
PR
     09-AUG-2002; 2002US-0402256P.
PR
     09-AUG-2002; 2002US-0402389P.
PR
PR
     12-AUG-2002; 2002US-0402786P.
PR
     12-AUG-2002; 2002US-0402816P.
PR
     12-AUG-2002; 2002US-0402821P.
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
PR
     13-AUG-2002; 2002US-0403459P.
PR
     13-AUG-2002; 2002US-0403531P.
PR
     13-AUG-2002; 2002US-0403532P.
PR
PR
     13-AUG-2002; 2002US-0403563P.
PR
     13-AUG-2002; 2002US-0406317P.
     15-AUG-2002; 2002US-0403617P.
PR
     26-AUG-2002; 2002US-0406182P.
PR
PR
     26-AUG-2002; 2002US-0406355P.
PR
     27-AUG-2002; 2002US-0406240P.
PR
     12-SEP-2002; 2002US-0410084P.
     20-SEP-2002; 2002US-0412528P.
PR
     23-SEP-2002; 2002US-0412731P.
PR
     30-SEP-2002; 2002US-0414801P.
PR
     30-SEP-2002; 2002US-0414839P.
PR
     30-SEP-2002; 2002US-0414840P.
PR
     30-SEP-2002; 2002US-0414954P.
PR
PR
     09-OCT-2002; 2002US-0417186P.
     09-OCT-2002; 2002US-0417406P.
PR
     23-OCT-2002; 2002US-0420639P.
PR
     28-OCT-2002; 2002US-0421156P.
PR
     31-OCT-2002; 2002US-0422690P.
PR
PR
     01-NOV-2002; 2002US-0423130P.
PR
     05-NOV-2002; 2002US-00423798.
PR
     05-NOV-2002; 2002US-0423798P.
     12-NOV-2002; 2002US-0425453P.
PR
XX
```

```
PA
     (CURA-) CURAGEN CORP.
XX
    Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PI
    Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PΙ
    Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PΙ
    Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PΙ
    Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
PΙ
    Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L; Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
ΡI
ΡI
ΡI
    Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
    Zhong H;
PΙ
XX
    WPI; 2004-081935/08.
DR
    P-PSDB; ADH71628.
DR
XX
PT
    New NOVX polypeptides and nucleic acid molecules useful for preventing or
    treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
PT
    obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS
    Example 21; SEQ ID NO 523; 1880pp; English.
XX
CC
    The invention relates to a novel isolated polypeptide (NOVX). A
CC
    polypeptide of the invention has cytostatic, immunomodulator,
    neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
    antilipaemic activity, and may have a use in gene therapy, and as a
CC
CC
    vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC
    any of the 303 fully defined nucleotide sequences given in the
    specification. The polypeptide is useful in the manufacture of a
CC
CC
    medicament for treating a syndrome associated with a human disease. The
    polypeptide, polynucleotide and antibody are useful in diagnosing,
CC
CC
    treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
    Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC
    diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC
    further used as hybridisation probes, in chromosome mapping, tissue
CC
CC
    typing, preventive medicine, and pharmacogenomics. The present sequence
CC
    encodes a NOVX polypeptide of the invention.
XX
    Sequence 2881 BP; 525 A; 985 C; 869 G; 502 T; 0 U; 0 Other;
SQ
                        97.2%;
                               Score 2674.8; DB 12; Length 2881;
 Query Match
 Best Local Similarity
                        98.9%;
                               Pred. No. 0;
 Matches 2727; Conservative
                              0:
                                 Mismatches
                                              22:
                                                   Indels.
                                                                Gaps
                                                                        3;
           1 CCGCGGGGCCCGCCCGCCCGCCCGCCGCCGCCGCCATGGCCGTCCGGCCC 60
Qу
             Db
          42 CCGCGGGCCCCGCCCCGCCCGCCCGCCCGCCGCGCCATGGCCGTCCGGCCC 101
          61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120
Qy
             102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161
Db
         121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180
Qу
             162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221
Db
Qу
         181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240
```

| Db   | 222  | TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC         | 281  |
|------|------|--|------|
| Qу   | 241  | AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG         | 300  |
| Db   | 282  | AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG         | 341  |
| Qy   | 301  | GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC         | 360  |
| Db   | 342  | GACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGTGAGCCGACCATGGAGGTCCGC         | 401  |
| Qy   | 361  | ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG         | 420  |
| Db   | 402  | ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG         | 461  |
| Qу   | 421  | TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC         | 480  |
| Db   | 462  | TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC         | 521  |
| Qy   | 481  | AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC         | 540  |
| Db   | 522  | AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC         | 581  |
| Qу   | 541  | ATCGTGCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG              | 600  |
| Db   | 582  | ATCGTGCTGCCGTCCACCGGAGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG               | 641  |
| Qу   | 601  | AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC         | 660  |
| Db   | 642  | AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC         | 701  |
| Qу   | 661  | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC         | 720  |
| Db   | 702  |  | 761  |
| Qу   | 721  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG             | 780  |
| Db   | 762  | ATCGTGGCACGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG                | 821  |
| Qу   | 781  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG         | 840  |
| Db   | 822  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG         | 881  |
| Qy · | 841  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT         | 900  |
| Db   | 882  | ${\tt AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT}$ | 941  |
| Qy   | 901  | GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG            | 957  |
| Db   | 942  | GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG         | 1001 |
| Qy   | 958  | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC         | 1017 |
| Db   | 1002 | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC         | 1061 |
| Qy   | 1018 | CCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT           | 1077 |
| Db   | 1062 | CCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT           | 1121 |

| Qу   | 1078 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC | 1137  |
|------|------|--|-------|
| Db   | 1122 | ACCAGTGACCTCTGTGTGCACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC | 1181  |
| Qу   | 1138 | CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTCTCATCCTCGTTTATTGC     | 1197  |
| Db   | 1182 | CTCATCGCCGTGGCCTGCTGCTGCTGCTGCTGCTCATCCTCGTTTATTGC           | 1241  |
| Qу   | 1198 | CGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC  | 1257  |
| Db   | 1242 | CGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC  | 1,301 |
| Qу   | 1258 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG | 1317  |
| Db   | 1302 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG | 1361  |
| Qy   | 1318 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC    | 1377  |
| Db   | 1362 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC    | 1418  |
| Qу   | 1378 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC | 1437  |
| Db   | 1419 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC | 1478  |
| Qу   | 1438 | ACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC | 1497  |
| Db   | 1479 | ACACTGCACCACACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC          | 1538  |
| Qy   | 1498 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC | 1557  |
| Db   | 1539 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC | 1598  |
| Qу   | 1558 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT  | 1617  |
| Db · | 1599 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT  | 1658  |
| Qу   | 1618 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG | 1677  |
| Db   | 1659 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG | 1718  |
| QУ   | 1678 | AGGTTGCCCCTÁGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT | 1737  |
| Db   | 1719 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT | 1778  |
| Qу   | 1738 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1797  |
| Dþ   | 1779 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1838  |
| Qу   | 1798 | GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTG    | 1854  |
| Db   | 1839 | GACAGCTGGAGCCTCCAAAAAGCAGTCGTGCGAGGCAGCTGGGAGCAGGATGTG       | 1898  |
| Qу   | 1855 | CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC | 1914  |
| Db   | 1899 | CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC | 1958  |

| Qy |       | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG |      |
|----|-------|--|------|
| Db |       | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG |      |
| Qy | 1975  | GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC    | 2034 |
| Db | 2019  | GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC    | 2078 |
| Qy | 2035  | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG | 2094 |
| Db | 2079  | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG | 2138 |
| Qу | 2095  | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC | 2154 |
| Db | 2139  | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC | 2198 |
| Qу | 2155  | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG | 2214 |
| Db | 2199  | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG | 2258 |
| Qу | 2215  | CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC | 2274 |
| Db | 2259  | CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC | 2318 |
| Qу | 2275  | TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG | 2334 |
| Db | 2319  | TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG | 2378 |
| Qу | 2335. | CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC | 2394 |
| Db | 2379  | CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC | 2438 |
| Qу | 2395  | AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG | 2454 |
| Db | 2439  | AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG | 2498 |
| Qу | 2455  | GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG | 2514 |
| Db | 2499  | GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG | 2558 |
| Qy | 2515  | GACCCACCCTGTAGGCGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC  | 2574 |
| Db | 2559  |  | 2618 |
| Qу | 2575  | AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCACAGCCATGATCCTCAACCTGTGG  | 2634 |
| Db | 2619  |  | 2678 |
| Qу | 2635  | GAGGCGCGCACTTCCCCAACGCCAACCTCAGCCAGCTGGCTG                   | 2694 |
| Db | 2679  | GAGGCGCGGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTG                  | 2738 |
| Qу | 2695  | GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCC     | 752  |
| Db | 2739  |  | 796  |

```
RESULT 14
ADH71639
     ADH71639 standard; DNA; 2881 BP.
ID
XX
AC
     ADH71639;
XX
DT
     25-MAR-2004
                  (first entry)
XX
DE
     Human gene of the invention NOV21p SEQ ID NO:535.
XX
KW
     ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
KW
     anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
     vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;
KW
     obesity; diabetes; infectious disease; metabolic syndrome X;
KW
     dyslipidaemia.
XX
os
     Homo sapiens.
XX
PN
     WO2003102155-A2.
XX
PD
     11-DEC-2003.
XX
PF
     03-JUN-2003; 2003WO-US017430.
XX
PR
     03-JUN-2002; 2002US-0385120P.
PR
     04-JUN-2002; 2002US-0385784P.
PR
     05-JUN-2002; 2002US-0386041P.
PR
     05-JUN-2002; 2002US-0386047P.
     06-JUN-2002; 2002US-0386376P.
PR
     06-JUN-2002; 2002US-0386453P.
PR
PR
     06-JUN-2002; 2002US-0386864P.
PR
     06-JUN-2002; 2002US-0387016P.
PR
     07-JUN-2002; 2002US-0386796P.
PR
     07-JUN-2002; 2002US-0386816P.
     07-JUN-2002; 2002US-0386931P.
PR
PR
     07-JUN-2002; 2002US-0386942P.
     07-JUN-2002; 2002US-0386971P.
PR
     07-JUN-2002; 2002US-0387262P.
PR
     08-JUN-2002; 2002US-0296960P.
PR
PR
     10-JUN-2002; 2002US-0387400P.
PR
     10-JUN-2002; 2002US-0387535P.
PR
     11-JUN-2002; 2002US-0387610P.
     11-JUN-2002; 2002US-0387625P.
PR
     11-JUN-2002; 2002US-0387634P.
PR
     11-JUN-2002; 2002US-0387668P.
PR
PR
     11-JUN-2002; 2002US-0387696P.
PR
     11-JUN-2002; 2002US-0387702P.
PR
     11-JUN-2002; 2002US-0387836P.
PR
     11-JUN-2002; 2002US-0387859P.
     12-JUN-2002; 2002US-0387933P.
PR
     12-JUN-2002; 2002US-0387934P.
PR
     12-JUN-2002; 2002US-0387960P.
PR
PR
     12-JUN-2002; 2002US-0388022P.
PR
     12-JUN-2002; 2002US-0388096P.
PR
     13-JUN-2002; 2002US-0389123P.
PR
     14-JUN-2002; 2002US-0389118P.
```

14-JUN-2002; 2002US-0389120P.

PR

```
14-JUN-2002; 2002US-0389144P.
PR
     14-JUN-2002; 2002US-0389146P.
PR
     17-JUN-2002; 2002US-0389729P.
PR
PR
     17-JUN-2002; 2002US-0389742P.
     18-JUN-2002; 2002US-0389884P.
PR
     19-JUN-2002; 2002US-0390006P.
PR
     19-JUN-2002; 2002US-0390209P.
PR
     21-JUN-2002; 2002US-0390763P.
PR
     17-JUL-2002; 2002US-0396706P.
PR
     06-AUG-2002; 2002US-0401628P.
PR
     09-AUG-2002; 2002US-0402156P.
PR
PR
     09-AUG-2002; 2002US-0402256P.
     09-AUG-2002; 2002US-0402389P.
PR
     12-AUG-2002; 2002US-0402786P.
PR
     12-AUG-2002; 2002US-0402816P.
PR
     12-AUG-2002; 2002US-0402821P.
PR
PR
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
PR
     13-AUG-2002; 2002US-0403459P.
     13-AUG-2002; 2002US-0403531P.
PR
     13-AUG-2002; 2002US-0403532P.
PR
     13-AUG-2002; 2002US-0403563P.
PR
     13-AUG-2002; 2002US-0406317P.
PR
     15-AUG-2002; 2002US-0403617P.
PR
     26-AUG-2002; 2002US-0406182P.
PR
PR
     26-AUG-2002; 2002US-0406355P.
PR
     27-AUG-2002; 2002US-0406240P.
     12-SEP-2002; 2002US-0410084P.
PR
     20-SEP-2002; 2002US-0412528P.
PR
     23-SEP-2002; 2002US-0412731P.
PR
     30-SEP-2002; 2002US-0414801P.
PR
     30-SEP-2002; 2002US-0414839P.
PR
PR
     30-SEP-2002; 2002US-0414840P.
     30-SEP-2002; 2002US-0414954P.
PR
     09-OCT-2002; 2002US-0417186P.
PR
     09-OCT-2002; 2002US-0417406P.
PR
     23-OCT-2002; 2002US-0420639P.
PR
     28-OCT-2002; 2002US-0421156P.
PR
     31-OCT-2002; 2002US-0422690P.
PR
PR
     01-NOV-2002; 2002US-0423130P.
     05-NOV-2002; 2002US-00423798.
PR
PR
     05-NOV-2002; 2002US-0423798P.
     12-NOV-2002; 2002US-0425453P.
PR
XX
PA
     (CURA-) CURAGEN CORP.
XX
     Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
PΙ
PΙ
     Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
ΡI
     Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;
PΙ
     Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PΙ
     Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
     Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D, Ras
Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;
PΙ
                                                                    Rastelli L;
PΙ
     Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
PΙ
PΙ
     Zhong H;
XX
DR
     WPI; 2004-081935/08.
```

```
DR
    P-PSDB; ADH71640.
XX
    New NOVX polypeptides and nucleic acid molecules useful for preventing or
РΤ
    treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
    obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
PT
XX
PS
    Example 21; SEQ ID NO 535; 1880pp; English.
XX
CC
    The invention relates to a novel isolated polypeptide (NOVX). A
CC
    polypeptide of the invention has cytostatic, immunomodulator,
    neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
    antilipaemic activity, and may have a use in gene therapy, and as a
CC
CC
    vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC
    any of the 303 fully defined nucleotide sequences given in the
    specification. The polypeptide is useful in the manufacture of a
CC
CC
    medicament for treating a syndrome associated with a human disease. The
    polypeptide, polynucleotide and antibody are useful in diagnosing,
CC
    treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC
    Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC
    diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC
CC
    further used as hybridisation probes, in chromosome mapping, tissue
    typing, preventive medicine, and pharmacogenomics. The present sequence
CC
    encodes a NOVX polypeptide of the invention.
CC
XX
    Sequence 2881 BP; 525 A; 985 C; 869 G; 502 T; 0 U; 0 Other;
SQ
 Query Match
                     97.2%; Score 2674.8; DB 12; Length 2881;
 Best Local Similarity
                     98.9%; Pred. No. 0;
 Matches 2727; Conservative
                           0: Mismatches
                                          22;
                                              Indels
                                                       9;
                                                          Gaps
                                                                 3;
          Qу
           42 CCGCGGGGCCCGGCCCGCCCGCCTGCCCGCCGCGCCATGGCCGTCCGGCCC 101
Db
         61 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 120
Qу
           102 GGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTCCGCGGCTCGGGTGCC 161
Db
        121 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 180
Qу
            162 CAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCGGACCTGCTTCCCCAC 221
Db
        181 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 240
Qy
            222 TTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCAGTGCTGCTTGTGTGC 281
Db
        241 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 300
Qу
            282 AAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAGTGGGTGCGCCAGGTG 341
Db
        301 GACCACGTGATCGAGCGCAGCAGACGGGAGCAGTGGGCTGCCCACCATGGAGGTCCGC 360
Qу
            111 111111111111
        342 GACCACGTGATCGAGCGCAGCAGACGGGGGGGGGGGGTGGTGAGCCGACCATGGAGGTCCGC 401
Db
        361 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 420
Qу
```

402 ATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAGGAATACTGGTGCCAG 461

Db

| QУ   | 421  | TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC | 480  |
|------|------|--|------|
| Db   | 462  | TGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCCTACATCCGCATAGCC | 521  |
| Qy   | 481  | AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC | 540  |
| Db   | 522  | AGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTGTCCCTGGAGCAGGGC | 581  |
| Qy . | 541  | ATCGTGCTGCCCTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG | 600  |
| Db   | 582  | ATCGTGCTGCCCTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGG      | 641  |
| Qу   | 601  | AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC | 660  |
| Db   | 642  | AACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC     | 701  |
| Qy   | 661  | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC | 720  |
| Db   | 702  | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC | 761  |
| Qy   | 721  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG     | 780  |
| Db   | 762  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG     | 821  |
| Qy   | 781  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG | 840  |
| Db   | 822  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG | 881  |
| Qy   | 841  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT | 900  |
| Db   | 882  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT | 941  |
| Qу   | 901  | GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG    | 957  |
| Db   | 942  | GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG | 1001 |
| Qу   | 958  | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC | 1017 |
| Db   | 1002 | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC | 1061 |
| Qу   | 1018 | CCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT  | 1077 |
| Db   | 1062 | CCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT   | 1121 |
| Qу   | 1078 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC | 1137 |
| Db   | 1122 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC | 1181 |
| Qy   | 1138 | CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTGCTCATCCTCGTTTATTGC    | 1197 |
| Db   | 1182 | CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTCATCCTCGTTTATTGC       | 1241 |
| Qy   | 1198 | CGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC  | 1257 |
| Db   | 1242 | CGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC  | 1301 |

•

| Qу   | 1258 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG | 1317 |
|------|------|--|------|
| Db   | 1302 |  | 1361 |
| Qу   | 1318 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC    | 1377 |
| Db   | 1362 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC    | 1418 |
| Qy   | 1378 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC | 1437 |
| Db   | 1419 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC | 1478 |
| Qу   | 1438 | ACACTGCACCACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC  | 1497 |
| Db   | 1479 | ACACTGCACCACACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC          | 1538 |
| Qу   | 1498 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC | 1557 |
| Db   | 1539 |  | 1598 |
| Qу   | 1558 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT  | 1617 |
| Dp . | 1599 |  | 1658 |
| Qу   | 1618 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG | 1677 |
| Db   | 1659 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG | 1718 |
| Qу   | 1678 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT | 1737 |
| Db   | 1719 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT | 1778 |
| Qу   | 1738 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1797 |
| Db   | 1779 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1838 |
| Qу   | 1798 | GACAGCTGGAGCCTCCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTG         | 1854 |
| Db   | 1839 | GACAGCTGGAGCCTCAAAAAGCAGTCGTGCGAGGCAGCTGGGAGCAGGATGTG        | 1898 |
| Qу   | 1855 | CTGCACCTGGGCGAGGAGGCCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC  | 1914 |
| Db   | 1899 | CTGCACCTGGGCGAGGAGGCCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC  | 1958 |
| Qy   | 1915 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG | 1974 |
| Db   | 1959 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG | 2018 |
| Qy   | 1975 | GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC    | 2034 |
| Db   | 2019 | GCTGCCGCCAAGCGCCTCAGGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC    | 2078 |
| Qy   | 2035 | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG | 2094 |
| Db   | 2079 | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG | 2138 |
| Ωv   | 2095 | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC | 2154 |

```
2139 CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC 2198
Db
     2155 AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG 2214
Qy
         2199 AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG 2258
Db
     2215 CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC 2274
Qу
        2259 CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC 2318
Db
     2275 TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG 2334
Qу
        2319 TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG 2378
Db
     2335 CTGTGGGTGTGGCAGGTGGAGGGCGACGGCCAGACTTCAACATTCAACATCACC 2394
Qy
         2379 CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACATCACC 2438
Db
     2395 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2454
Qу
        2439 AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG 2498
Db
     2455 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2514
Qу
         2499 GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG 2558
Db
     2515 GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2574
Qy
         2559 GACCCACCTGTAGGCGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC 2618
Db
     2575 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2634
Qу
         2619 AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG 2678
Db
     Qy
        Db
     Qy
         Db
RESULT 15
ADH71643
   ADH71643 standard; DNA; 2881 BP.
ID
XX
AC
   ADH71643;
XX
DT
   25-MAR-2004 (first entry)
XX
DE
   Human gene of the invention NOV21r SEQ ID NO:539.
XX
KW
   ds; gene; human; cytostatic; immunomodulator; neuroprotective; nootropic;
   anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;
KW
```

vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

KW

```
obesity; diabetes; infectious disease; metabolic syndrome X;
KW
KW
     dyslipidaemia.
XX
     Homo sapiens.
OS
XX
     WO2003102155-A2.
PN
XX
PD
     11-DEC-2003.
XX
     03-JUN-2003; 2003WO-US017430.
PF
XX
PR
     03-JUN-2002; 2002US-0385120P.
PR
     04-JUN-2002; 2002US-0385784P.
     05-JUN-2002; 2002US-0386041P.
PR
     05-JUN-2002; 2002US-0386047P.
PR
     06-JUN-2002; 2002US-0386376P.
PR
     06-JUN-2002; 2002US-0386453P.
PR
PR
     06-JUN-2002; 2002US-0386864P.
PR
     06-JUN-2002; 2002US-0387016P.
PR
     07-JUN-2002; 2002US-0386796P.
     07-JUN-2002; 2002US-0386816P.
PR
     07-JUN-2002; 2002US-0386931P.
PR
     07-JUN-2002; 2002US-0386942P.
PR
PR
     07-JUN-2002; 2002US-0386971P.
PR
     07-JUN-2002; 2002US-0387262P.
PR
     08-JUN-2002; 2002US-0296960P.
     10-JUN-2002; 2002US-0387400P.
PR
     10-JUN-2002; 2002US-0387535P.
PR
     11-JUN-2002; 2002US-0387610P.
PR
     11-JUN-2002; 2002US-0387625P.
PR
     11-JUN-2002; 2002US-0387634P.
PR
     11-JUN-2002; 2002US-0387668P.
PR
PR
     11-JUN-2002; 2002US-0387696P.
PR
     11-JUN-2002; 2002US-0387702P.
     11-JUN-2002; 2002US-0387836P.
PR
     11-JUN-2002; 2002US-0387859P.
PR
     12-JUN-2002; 2002US-0387933P.
PR
PR
     12-JUN-2002; 2002US-0387934P.
PR
     12-JUN-2002; 2002US-0387960P.
     12-JUN-2002; 2002US-0388022P.
PR
PR
     12-JUN-2002; 2002US-0388096P.
     13-JUN-2002; 2002US-0389123P.
PR
     14-JUN-2002; 2002US-0389118P.
PR
     14-JUN-2002; 2002US-0389120P.
PR
     14-JUN-2002; 2002US-0389144P.
PR
PR
     14-JUN-2002; 2002US-0389146P.
PR
     17-JUN-2002; 2002US-0389729P.
PR
     17-JUN-2002; 2002US-0389742P.
PR
     18-JUN-2002; 2002US-0389884P.
PR
     19-JUN-2002; 2002US-0390006P.
     19-JUN-2002; 2002US-0390209P.
PR
PR
     21-JUN-2002; 2002US-0390763P.
PR
     17-JUL-2002; 2002US-0396706P.
PR
     06-AUG-2002; 2002US-0401628P.
PR
     09-AUG-2002; 2002US-0402156P.
PR
     09-AUG-2002; 2002US-0402256P.
PR
     09-AUG-2002; 2002US-0402389P.
```

```
12-AUG-2002; 2002US-0402786P.
PR
     12-AUG-2002; 2002US-0402816P.
PR
PR
     12-AUG-2002; 2002US-0402821P.
     12-AUG-2002; 2002US-0402832P.
PR
     13-AUG-2002; 2002US-0403448P.
PR
PR
    13-AUG-2002; 2002US-0403459P.
    13-AUG-2002; 2002US-0403531P.
PR
    13-AUG-2002; 2002US-0403532P.
PR
    13-AUG-2002; 2002US-0403563P.
PR
PR
     13-AUG-2002; 2002US-0406317P.
     15-AUG-2002; 2002US-0403617P.
PR
     26-AUG-2002; 2002US-0406182P.
PR
     26-AUG-2002; 2002US-0406355P.
PR
    27-AUG-2002; 2002US-0406240P.
PR
     12-SEP-2002; 2002US-0410084P.
PR
     20-SEP-2002; 2002US-0412528P.
PR
     23-SEP-2002; 2002US-0412731P.
PR
     30-SEP-2002; 2002US-0414801P.
PR
PR
     30-SEP-2002; 2002US-0414839P.
     30-SEP-2002; 2002US-0414840P.
PR
     30-SEP-2002; 2002US-0414954P.
PR
     09-OCT-2002; 2002US-0417186P.
PR
     09-OCT-2002; 2002US-0417406P.
PR
    23-OCT-2002; 2002US-0420639P.
PR
    28-OCT-2002; 2002US-0421156P.
PR
PR
     31-OCT-2002; 2002US-0422690P.
PR
     01-NOV-2002; 2002US-0423130P.
PR
     05-NOV-2002; 2002US-00423798.
     05-NOV-2002; 2002US-0423798P.
PR
PR
     12-NOV-2002; 2002US-0425453P.
XX
PA
     (CURA-) CURAGEN CORP.
XX
PΙ
    Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;
    Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;
PΙ
    Ettenberg S, Gangolli EA,
                                Gerlach VL, Gorman L, Gunther E, Guo X;
PΙ
     Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;
PI
PΙ
    Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;
     Padigaru M, Patturajan M, Pena CEA, Peyman JA, Raha D,
PΙ
                                                                 Rastelli L;
PI
     Rieger DK, Rothenberg ME,
                                 Sciore P, Shenoy SG,
                                                        Shimkets RA;
PΙ
                 Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;
     Smithson G,
PΙ
     Zhong H;
XX
DR
    WPI; 2004-081935/08.
DR
     P-PSDB; ADH71644.
XX
PT
    New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT
     treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT
     obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS
     Example 21; SEQ ID NO 539; 1880pp; English.
XX
CC
     The invention relates to a novel isolated polypeptide (NOVX). A
CC
     polypeptide of the invention has cytostatic, immunomodulator,
CC
     neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC
     antilipaemic activity, and may have a use in gene therapy, and as a
CC
     vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
```

CC any of the 303 fully defined nucleotide sequences given in the specification. The polypeptide is useful in the manufacture of a CC CC medicament for treating a syndrome associated with a human disease. The CC polypeptide, polynucleotide and antibody are useful in diagnosing, CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia, CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are CC CC further used as hybridisation probes, in chromosome mapping, tissue CC typing, preventive medicine, and pharmacogenomics. The present sequence CC encodes a NOVX polypeptide of the invention. XX

SQ Sequence 2881 BP; 527 A; 985 C; 867 G; 502 T; 0 U; 0 Other;

Query Match 97.2%; Score 2674.8; DB 12; Length 2881; Best Local Similarity 98.9%; Pred. No. 0; Matches 2727; Conservative 0; Mismatches 22; Indels 9; Gaps 3;

| Matches | 212 | ; Conservative      | υ;   | Mismatches   | 22;      | indels    | 9;    | Gaps   | 3;  |
|---------|-----|---------------------|------|--------------|----------|-----------|-------|--------|-----|
| Qy      | 1   | CCGCGGGGCCCGCGCCCG  |      |              |          |           |       |        | 60  |
| Db      | 42  | CCGCGGGGCCCG        |      |              |          |           |       |        | 101 |
| Qy      | 61  | GGCCTGTGGCCAGCGCTCC |      |              |          |           |       |        | 120 |
| Db      | 102 | GGCCTGTGGCCAGCGCTCC |      |              |          |           |       |        | 161 |
| Qy      | 121 | CAGCAGAGTGCCACCGTGG |      |              |          |           |       |        | 180 |
| Db      | 162 | CAGCAGAGTGCCACCGTGG |      |              |          |           |       |        | 221 |
| Qy      | 181 | TTCCTGGTGGAGCCCGAGG |      |              |          |           |       |        | 240 |
| Db      | 222 | TTCCTGGTGGAGCCCGAGG | ATG' | TGTACATCGTCA | AAGAACA. | AGCCAGTGC | TGCTT | GTGTGC | 281 |
| Qy      | 241 | AAGGCCGTGCCCGCCACGC |      |              |          |           |       |        | 300 |
| Db      | 282 | AAGGCCGTGCCCGCCACGC |      |              |          |           |       |        | 341 |
| Qy      | 301 | GACCACGTGATCGAGCGCA |      |              |          | TGCCCACCA |       |        | 360 |
| Db      | 342 | GACCACGTGATCGAGCGCA |      |              |          |           |       |        | 401 |
| Qу      | 361 | ATTAATGTCTCAAGGCAGC |      |              |          |           |       |        | 420 |
| Db      | 402 | ATTAATGTCTCAAGGCAGC |      |              |          |           |       |        | 461 |
| Qу      | 421 | TGCGTGGCATGGAGCTCCT |      |              |          |           |       |        | 480 |
| Db      | 462 | TGCGTGGCATGGAGCTCCT |      |              |          |           |       |        | 521 |
| Qу      | 481 | AGATTGCGCAAGAACTTCG |      |              |          |           |       |        | 540 |
| Db      | 522 | AGATTGCGCAAGAACTTCG |      |              |          |           |       |        | 581 |
| Qy      | 541 | ATCGTGCTGCCCTGCCGTC |      |              |          |           |       |        | 600 |
| Db      | 582 | ATCGTGCTGCCCTGCCGTC |      |              |          |           |       |        | 641 |

| Qу   | 601  | AACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC | 0,00 |
|------|------|--|------|
| Db   | 642  | AACGAGGACCTGGTGGACCCCTGGACCCCAATGTATACATCACGCGGGAGCACAGC     | 701  |
| Qу   | 661  | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC | 720  |
| Db   | 702  | CTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACCTGCGTGGCCAAGAAC | 761  |
| Qу   | 721  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG     | 780  |
| Db   | 762  | ATCGTGGCACGTCGCCGCAGCGCCTCCGCTGTCATCGTCTACGTGAACGGTGGGTG     | 821  |
| Qу   | 781  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG | 840  |
| Db   | 822  | TCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGCGGCTGGCAGAAACGG | 881  |
| Qу   | 841  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT | 900  |
| Db   | 882  | AGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTCTGTGAGGGGCAGAAT | 941  |
| Qу   | 901  | GTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGCTGGAGCCCGTGG    | 957  |
| Db   | 942  | GTCCATGACCGCACCGTCTCCTCTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGG | 1001 |
| Qу   | 958  | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC | 1017 |
| Db   | 1002 | AGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGTGAGTGCTCTGAC | 1061 |
| Qу   | 1018 | CCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT   | 1077 |
| Db   | 1062 | CCAGCACCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT   | 1121 |
| Qу   | 1078 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC | 1137 |
| Db   | 1122 | ACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC | 1181 |
| Qy   | 1138 | CTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTGCTCATCCTCGTTTATTGC    | 1197 |
| Db   | 1182 | CTCATCGCCGTGGCCTGCTGCTGCTGCTGCTGCTCATCCTCGTTTATTGC           | 1241 |
| Qу   | 1198 | CGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC  | 1257 |
| Db   | 1242 | CGGAAGAAGGAGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTCACCTCAGGCTTC  | 1301 |
| Qу   | 1258 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG | 1317 |
| Db · | 1302 | CAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTCACCATCCAGCCG | 1361 |
| Qу   | 1318 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC    | 1377 |
| Db   | 1362 | GACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGGCAGGATGGGCCC    | 1418 |
| Qу   | 1378 | AGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGTGGCGGCCGCCAC | 1437 |
| Db   | 1419 |  | 1478 |
| Ov   | 1438 | ACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCCCGCCTCTCCACC | 1497 |

| Db   | 1479 |  | 1538 |
|------|------|--|------|
| Qy   | 1498 | CAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTATGGGACCTTCAAC | 1557 |
| Db   | 1539 |  | 1598 |
| Qy   | 1558 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCAGAT  | 1617 |
| Db   | 1599 | TTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTCATCCCCCCAGAT | 1658 |
| Qу   | 1618 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG | 1677 |
| Dp . | 1659 | GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGGAAGACGTG | 1718 |
| Qу   | 1678 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT | 1737 |
| Db   | 1719 | AGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCCT | 1778 |
| Qy   | 1738 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1797 |
| Db   | 1779 | GGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGGGAGCCCAGCCCT | 1838 |
| Qy   | 1798 | GACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGGATGTG    | 1854 |
| Db   | 1839 | GACAGCTGGAGCCTCCAAAAAGCAGTCGTGCGAGGGCAGCTGGGAGCAGGATGTG      | 1898 |
| Qу   | 1855 | CTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC | 1914 |
| Db   | 1899 | CTGCACCTGGGCGAGGAGGCCCCTCCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC  | 1958 |
| Qy   | 1915 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG | 1974 |
| Db   | 1959 | TGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCCCTCAGCGTG | 2018 |
| QУ   | 1975 | GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC    | 2034 |
| Db   | 2019 | GCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACCTCCCTC    | 2078 |
| Qy   |      | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG |      |
| Db   | 2079 | TACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAGGTGGTGCAG | 2138 |
| Qу   | 2095 | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC | 2154 |
| Db   |      | CTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCACTTCAAGGAC |      |
| Qу   | 2155 | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG | 2214 |
| Db   | 2199 | AGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGGAAGAGTAAG | 2258 |
| Qy   | 2215 | CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGGCACGCAGCGGTAC | 2274 |
| Db   | 2259 | CTCCTTGTCAGCTACCAGGAGATCCCCTTTTATCACATCTGGAATGACACGCAGCGGTAC | 2318 |
| Qу   | 2275 | TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG | 2334 |

| Db   | 2319 | TTGCACTGCACCTTCACCCTGGAGCGTGTCAGCCCCAGCACTAGTGACCTGGCCTGCAAG | 2378   |
|------|------|--|--------|
| Qу   | 2335 | CTGTGGGTGTGGCAGGTGGAGGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC | 2394   |
| Db   | 2379 | CTGTGGGTGTGGCAGGTGGAGGCGACGGGCAGAGCTTCAGCATCAACTTCAACATCACC  | 2438   |
| Qy   | 2395 | AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG | 2454   |
| Db   | 2439 | AAGGACACAAGGTTTGCTGAGCTGCTGGCTCTGGAGAGTGAAGCGGGGGTCCCAGCCCTG | 2498   |
| Qу   | 2455 | GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG | 2514   |
| Db   | 2499 | GTGGGCCCCAGTGCCTTCAAGATCCCCTTCCTCATTCGGCAGAAGATAATTTCCAGCCTG | 2558   |
| Qу   | 2515 | GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC | 2574   |
| Db   | 2559 | GACCCACCCTGTAGGCGGGGTGCCGACTGGCGGACTCTGGCCCAGAAACTCCACCTGGAC | 2618   |
| Qу   | 2575 |  | 2634   |
| Db   | 2619 | AGCCATCTCAGCTTCTTTGCCTCCAAGCCCAGCCCCACAGCCATGATCCTCAACCTGTGG | 2678   |
| Qy   | 2635 |  | 2694   |
| Db   | 2679 | GAGGCGCGCACTTCCCCAACGGCAACCTCAGCCAGCTGGCTG                   | 2738 - |
| Qy . | 2695 | GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCC     | 52     |
| Db   | 2739 | GGCCAGCCAGACGCTGGCCTCTTCACAGTGTCGGAGGCTGAGTGCTGAGGCCGGCC     | 96     |

Search completed: March 5, 2005, 22:58:50 Job time: 1411 secs

## GenCore version 5.1.6 Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 5, 2005, 19:26:23; Search time 433.33 Seconds

(without alignments)

10391.692 Million cell updates/sec

Title: US-10-624-932-1

Perfect score: 2752

Sequence: 1 ccgcggggccccgcgcccgg.....tqagtgctgaggccqgccag 2752

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*

2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*

3: /cgn2 6/ptodata/1/ina/6A COMB.seq:\*

4: /cgn2 6/ptodata/1/ina/6B COMB.seq:\*

5: /cgn2 6/ptodata/1/ina/PCTUS COMB.seq:\*

6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

|        |        | ₹     |        |              |                    | •                 |
|--------|--------|-------|--------|--------------|--------------------|-------------------|
| Result |        | Query |        |              |                    |                   |
| No.    | Score  | Match | Length | DB           | ID                 | Description       |
| 1      | 2259   | 82.1  | 3014   | <del>-</del> | US-08-808-982-1    | Sequence 1, Appli |
| 2      | 2259   | 82.1  | 3014   | 3            | US-09-306-902A-1   | Sequence 1, Appli |
| 3      | 1562.4 | 56.8  | 1787   | 2            | US-08-808-982-2    | Sequence 2, Appli |
| 4      | 1562.4 | 56.8  | 1787   | 3            | US-09-306-902A-2   | Sequence 2, Appli |
| 5      | 841.4  | 30.6  | 2831   | 2            | US-08-808-982-3    | Sequence 3, Appli |
| 6      | 841.4  | 30.6  | 2831   | 3            | US-09-306-902A-3   | Sequence 3, Appli |
| 7      | 833.6  | 30.3  | 3008   | 4            | US-09-949-016-4794 | Sequence 4794, Ap |
| 8      | 487    | 17.7  | 2736   | 4            | US-09-969-532-9    | Sequence 9, Appli |
| 9      | 487    | 17.7  | 3411   | 4            | US-09-969-532-33   | Sequence 33, Appl |
| 10     | 467    | 17.0  | 2703   | 4            | US-09-969-532-11   | Sequence 11, Appl |
| 11     | 460.2  | 16.7  | 2694   | 4            | US-09-969-532-13   | Sequence 13, Appl |
|        |        |       |        |              |                    |                   |

```
439.8
                16.0
                       2661 4
   12
                               US-09-969-532-15
                                                           Sequence 15, Appl
                       349 4
   13
          323
                11.7
                               US-09-471-276-345
                                                           Sequence 345, App
   14
        293.2
                10.7
                       1968 4
                               US-09-969-532-31
                                                           Sequence 31, Appl
   15
        293.2
                10.7
                      2001 4
                               US-09-969-532-29
                                                           Sequence 29, Appl
   16
        293.2
                10.7
                      2010 4
                               US-09-969-532-27
                                                           Sequence 27, Appl
   17
        293.2
                10.7
                      2043 4
                               US-09-969-532-25
                                                           Sequence 25, Appl
   18
                                US-09-969-532-7
        284.8
                10.3
                      1659 4
                                                           Sequence 7, Appli
   19
        284.8
                               US-09-969-532-5
                10.3
                      1692 · 4
                                                           Sequence 5, Appli
        284.8
   20
                10.3
                      1701 4
                               US-09-969-532-3
                                                           Sequence 3, Appli
                      1734 4
   21
        284.8
                10.3
                               US-09-969-532-1
                                                           Sequence 1, Appli
   22
          269
                 9.8
                       771 1
                                                           Sequence 17, Appl
С
                               US-08-253-155A-17
  23
        182.2
                 6.6 19326 4
                               US-09-949-016-16776
                                                           Sequence 16776, A
   24
        123.4
                 4.5 114139
                               US-09-949-016-16536
                                                           Sequence 16536, A
   25
         119
                 4.3
                       305 2
                               US-08-808-982-4
                                                           Sequence 4, Appli
   26
         119
                 4.3
                       305 3
                               US-09-306-902A-4
                                                           Sequence 4, Appli
  27
         80.4
                 2.9
                       601 4
С
                               US-09-949-016-169999
                                                           Sequence 169999,
   28
        80.2
                 2.9
                       966 4
                               US-09-969-532-23
                                                           Sequence 23, Appl
   29
        75.8
                 2.8
                      1008 4
                               US-09-969-532-19
                                                           Sequence 19, Appl
С
  30
        73.8
                2.7
                     601 4
                               US-09-949-016-170033
                                                           Sequence 170033,
      . 66.4
  31
                2.4
                      601 4
С
                               US-09-949-016-170040
                                                           Sequence 170040,
  32
        66.4
                 2.4
                      601 4
                               US-09-949-016-170041
                                                          Sequence 170041,
   33
         63.6
                2.3
                      5784 4
                               US-09-949-016-462
                                                          Sequence 462, App
                      601 4
С
  34
          63
                2.3
                               US-09-949-016-170000
                                                          Sequence 170000,
  35
          63
                 2.3
                               US-09-949-016-170034
                       601 4
                                                          Sequence 170034,
   36
         60.2
                 2.2
                      2820
                            4
                               US-09-854-845-15
                                                          Sequence 15, Appl
   37
         60.2
                2.2
                      2865 4
                               US-09-854-845-13
                                                          Sequence 13, Appl
   38
        60.2
                2.2
                      3105 4
                               US-09-854-845-5
                                                          Sequence 5, Appli
   39
        60.2
                2.2
                      3150 4
                               US-09-854-845-1
                                                          Sequence 1, Appli
   40
        60.2
                2.2
                      3237 4
                               US-09-854-845-7
                                                          Sequence 7, Appli
   41
        60.2
                2.2
                      3282 4
                               US-09-854-845-3
                                                          Sequence 3, Appli
   42
        60.2
                2.2
                      3411 4
                               US-09-854-845-11
                                                          Sequence 11, Appl
   43
        60.2
                2.2
                      3456 4
                               US-09-854-845-9
                                                          Sequence 9, Appli
   44
        60.2
                2.2
                      4074 4
                               US-09-854-845-17
                                                          Sequence 17, Appl
   45
        56.4
                2.0
                      5178 4
                               US-09-949-016-5241
                                                          Sequence 5241, Ap
```

## ALIGNMENTS

```
RESULT 1
US-08-808-982-1
; Sequence 1, Application US/08808982
; Patent No. 5939271
  GENERAL INFORMATION:
    APPLICANT: Tessier-Lavigne, Marc
    APPLICANT: Leonardo, E. David
    APPLICANT: Hink, Lindsay
    APPLICANT: Masu, Masayuki
    APPLICANT: Kazuko, Keino-Masu
    TITLE OF INVENTION: Netrin Receptors
    NUMBER OF SEQUENCES: 8
    CORRESPONDENCE ADDRESS:
      ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
;
      STREET: 268 BUSH STREET, SUITE 3200
      CITY: SAN FRANCISCO
      STATE: CALIFORNIA
;
      COUNTRY: USA
```

```
ZIP: 94104
   COMPUTER READABLE FORM:
     MEDIUM TYPE: Floppy disk
     COMPUTER: IBM PC compatible
     OPERATING SYSTEM: PC-DOS/MS-DOS
     SOFTWARE: PatentIn Release #1.0, Version #1.30
   CURRENT APPLICATION DATA:
     APPLICATION NUMBER: US/08/808,982
     FILING DATE:
     CLASSIFICATION: 530
   ATTORNEY/AGENT INFORMATION:
     NAME: OSMAN, RICHARD A
     REGISTRATION NUMBER: 36,627
     REFERENCE/DOCKET NUMBER: UC96-217
   TELECOMMUNICATION INFORMATION:
     TELEPHONE: (415) 343-4341
     TELEFAX: (415) 343-4342
  INFORMATION FOR SEQ ID NO: 1:
   SEQUENCE CHARACTERISTICS:
     LENGTH: 3014 base pairs
     TYPE: nucleic acid
     STRANDEDNESS: double
     TOPOLOGY: linear
   MOLECULE TYPE: cDNA
US-08-808-982-1
 Query Match
                    82.1%; Score 2259; DB 2; Length 3014;
 Best Local Similarity 89.7%; Pred. No. 0;
 Matches 2427; Conservative
                         0; Mismatches 280; Indels
                                                     Gaps
                                                  0;
                                                           0;
        46 ATGGCCGTCCGGCCCGGCCTGTGGCCAGCGCTCCTGGGCATAGTCCTCGCCGCTTGGCTC 105
Qу
          1 ATGGCCGTCCGGCCTGTGGCCAGTGCTCCTGGGCATAGTCCTCGCCGCCTGGCTT 60
Db
       106 CGCGGCTCGGGTGCCCAGCAGAGTGCCACCGTGGCCAACCCAGTGCCTGGTGCCAACCCG 165
Qy
          Db
        61 CGTGGTTCGGGTGCCCAGCAGAGTGCCACGGTGCCAATCCAGTGCCCGGTGCCAACCCC 120
       166 GACCTGCTTCCCCACTTCCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAAGCCA 225
Qу
          121 GACCTGCCCCCACTTCCTGGTAGAGCCTGAGGACGTGTACATTGTCAAGAACAAGCCG 180
Db
       226 GTGCTGCTTGTGCAAGGCCGTGCCCGCCACGCAGATCTTCTTCAAGTGCAACGGGGAG 285
Qу
          181 GTGTTGTTGGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTGCAATGGGGAA 240
Db
       286 TGGGTGCGCCAGGTGACCACGTGATCGAGCGCAGCACAGACGGGAGCAGTGGGCTGCCC 345
Qу
          Db
       346 ACCATGGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCGAGAAGGTGTTCGGGCTGGAG 405
Qу
          Db
       301 ACCATGGAGGTCCGTATCAACGTATCGAGGCAGCAGGTAGAGAAAGTGTTTGGGCTGGAG 360
       406 GAATACTGGTGCCAGTGCGTGGCATGGAGCTCCTCGGGCACCACCAAGAGTCAGAAGGCC 465
Qу
          Db
       361 GAATACTGGTGCCAGTGTGTGGCATGGAGCTCCTCGGGTACCACCAAAAGTCAGAAGGCC 420
```

| Qy   | 466  | TACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGCAGGAGCCGCTGGCCAAGGAGGTG | 525  |
|------|------|--|------|
| Db   | 421  | TACATCCGGATTGCCTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGCCAAGGAAGTG | 480  |
| Qу   | 526  | TCCCTGGAGCAGGGCATCGTGCCGTGCCGTCCACCGGAGGGCATCCCTCCAGCCGAG    | 585  |
| Db   | 481  | TCACTGGAGCAAGGCATTGTACTACCTTGTCGCCCCCAGAAGGAATCCCCCCAGCTGAG  | 540  |
| Qy   | 586  | GTGGAGTGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATACATC | 645  |
| Db   | 541  |  | 600  |
| Qу   | 646  | ACGCGGGAGCACAGCCTGGTGGTGCGACAGGCCCGCCTTGCTGACACGGCCAACTACACC | 705  |
| Db   | 601  | ACGCGGGAGCACAGCCTAGTCGTGCGTCAGGCCCGCCTGGCCGACACGGCCAACTACACC | 660  |
| Qу   | 706  | TGCGTGGCCAAGAACATCGTGGCACGTCGCCGCAGCGCCTCCGCTGCTGTCATCGTCTAC | 765  |
| Db   | 661  | TGTGTGGCCAAGAACATCGTAGCCCGTCGCCGAAGCACCTCTGCAGCGGTCATTGTTTAT | 720  |
| Qу   | 766  | GTGAACGGTGGTCGACGTGGACCGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGC     | 825  |
| Db   | 721  | GTGAACGGTGGTCGACGTGGACTGAGTGGTCCGTCTGCAGCGCCAGCTGTGGGCGT     | 780  |
| Qy   | 826  | GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCGCCTCTCAACGGGGGCGCTTTC | 885  |
| Db   | 781  | GGCTGGCAGAAACGGAGCCGGAGCTGCACCAACCCGGCACCTCTCAACGGGGGCGCCTTC | 840  |
| Qу   | 886  | TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACCCTGTGCCCAGTAGACGGCAGC | 945  |
| Db   | 841  | TGTGAGGGGCAGAATGTCCAGAAAACAGCCTGCGCCACTCTGTGCCCAGTGGATGGGAGC | 900  |
| QУ   | 946  | TGGAGCCCGTGGAGCAAGTGGTCGGCCTGTGGGCTGGACTGCACCCACTGGCGGAGCCGT | 1005 |
| Db   | 901  | TGGAGTTCGTGGAGTAAGTGGTCAGCCTGTGGGCTTGACTGCACCCACTGGCGGAGCCGC | 960  |
| Qy 1 | 1006 | GAGTGCTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTGCCAGGGCACTGACCTGGAC  | 1065 |
| Db   | 961  | GAGTGCTCTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTCGGGGTGCTGACCTGGAC | 1020 |
| Qy 1 | 1066 | ACCCGCAACTGTACCAGTGACCTCTGTGTACACAGTGCTTCTGGCCCTGAGGACGTGGCC | 1125 |
| Db 1 | 1021 | ACCCGCAACTGTACCAGTGACCTCTGCCTGCACACCGCTTCTTGCCCCGAGGACGTGGCT | 1080 |
| Qy 1 | 126  | CTCTATGTGGGCCTCATCGCCGTGGCCGTCTGCCTGGTCCTGCTGCTGCTCATC       | 1185 |
| Db 1 | 1081 | CTCTACATCGGCCTGTGGCTGTGTGCCTCTTCTTGCTGTTGCTGGCCCTTGGA        | 1140 |
| Qy 1 | 186  | CTCGTTTATTGCCGGAAGAAGGAGGGGCTGGACTCAGATGTGGCTGACTCGTCCATTCTC | 1245 |
| Db 1 | 141  | CTCATTTACTGTCGCAAGAAGGAAGGGCTGGACTCCGATGTGGCCGACTCGTCCATCCTC | 1200 |
| Qy 1 | 246  | ACCTCAGGCTTCCAGCCCGTCAGCATCAAGCCCAGCAAAGCAGACAACCCCCATCTGCTC | 1305 |
| Db 1 | 201  | ACCTCGGGCTTCCAGCCTGTCAGCATCAAGCCCAGCAAAGCAGACACCCCCACCTGCTC  | 1260 |

| QУ | 1306 | ACCATCCAGCCGGACCTCAGCACCACCACCACCTACCAGGGCAGTCTCTGTCCCCGG            | 1365 |
|----|------|--|------|
| Db | 1261 | ACCATCCAGCCAGACCTCAGCACCACCACCACCAGGGCAGTCTATGTTCGAGG                | 1320 |
| QУ | 1366 | CAGGATGGGCCCAGCCCCAAGTTCCAGCTCACCAATGGGCACCTGCTCAGCCCCCTGGGT         | 1425 |
| Db | 1321 | CAGGATGGACCCAGCCCCAAGTTCCAGCTCTCTAATGGTCACCTGCTCAGCCCACTGGGG         | 1380 |
| Qу | 1426 | GGCGGCCGCCACACTGCACCACAGCTCTCCCACCTCTGAGGCCGAGGAGTTCGTCTCC           | 1485 |
| Db | 1381 | AGTGGCCGCCATACGTTGCACCACAGCTCACCCACCTCTGAGGCTGAGGACTTCGTCTCC         | 1440 |
| Qу | 1486 | CGCCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGCACCAGCAACATGACCTAT         | 1545 |
| Db | 1441 | CGCCTCTCCACCCAAAACTACTTTCGTTCCCTGCCCCGCGGCACCAGCAACATGGCCTAC         | 1500 |
| Qу | 1546 | GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACAGGTATCAGCCTCCTC         | 1605 |
| Db | 1501 | GGGACCTTCAACTTCCTCGGGGGCCGGCTGATGATCCCTAATACGGGGATCAGCCTCCTC         | 1560 |
| Qу | 1606 | ATCCCCCAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAG          | 1665 |
| Db | 1561 | ATACCCCGGATGCCATCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCACAAG           | 1620 |
| QУ | 1666 | CCGGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCCATCGTTAGC         | 1725 |
| Db | 1621 | CCAGAAGACGTGAGGTTGCCCCTAGCTGGCTGTCAGACCCTGCTGAGTCCAGTCGTTAGC         | 1680 |
| Qy | 1726 | TGTGGACCCCTGGCGTCCTGCTCACCCGGCCAGTCATCCTGGCTATGGACCACTGTGGG          | 1785 |
| Db | 1681 | TGTGGGCCCCCAGGAGTCCTGCTCACCCGGCCAGTCATCCTTGCAATGGACCACTGTGGA         | 1740 |
| Qу | 1786 | GAGCCCAGCCCTGACAGCTGGAGCCTGCGCCTCAAAAAGCAGTCGTGCGAGGGCAGCTGG         | 1845 |
| Db | 1741 | GAGCCCAGCCCTGACAGCTGGAGTCTGCGCCTCAAAAAGCAGTCCTGCGAGGGCAGTTGG         | 1800 |
| Qу | 1846 | GAGGATGTGCTGCACCTGGGCGAGGAGGCGCCCTCCCACCTCTACTACTGCCAGCTGGAG         | 1905 |
| Db | 1801 | GAGGATGTGCTGCACCTTGGTGAGGAGTCACCTTCCCACCTCTACTACTGCCAGCTGGAG         | 1860 |
| Qу | 1906 | GCCAGTGCCTGCTACGTCTTCACCGAGCAGCTGGGCCGCTTTGCCCTGGTGGGAGAGGCC         | 1965 |
| Db | 1861 | GCCGGGGCCTGCTATGTCTTCACGGAGCAGCTGGGCCGCTTTGCCCTGGTAGGAGAGGCC         | 1920 |
| Qу | 1966 | CTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTGTTTGCGCCGGTGGCCTGCACC         | 2025 |
| Db | 1921 | CTCAGCGTGGCTGCCACCAAGCGCCTCAGGCTCCTTCTGTTTGCTCCCGTGGCCTGTACG         | 1980 |
| Qy | 2026 | TCCCTCGAGTACAACATCCGGGTCTACTGCCTGCATGACACCCACGATGCACTCAAGGAG         | 2085 |
| Db | 1981 | TCCCTTGAGTACAACATCCGAGTGTACTGCCTACACGACACCCACGACGCTCTCAAGGAG         | 2040 |
| Qу | 2086 | GTGGTGCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGCCACGGGTCCTGCAC         | 2145 |
| Db | 2041 | GTGGTGCAGCTGGAGAAGCAGCTAGGTGGACAGCTGATCCAGGAGCCTCGCGTCCTGCAC         | 2100 |
| Qу | 2146 | ${\tt TTCAAGGACAGTTACCACAACCTGCGCCTATCCATCCACGATGTGCCCAGCTCCCTGTGG}$ | 2205 |